Sequences in Pairing Problems[☆]

A new approach to reconcile stability with strategy-proofness for elementary matching problems

JOB-MARKET PAPER

Jens Gudmundsson

Department of Economics, Lund University, Box 7082, SE-222 07 Lund, Sweden

Abstract

We study two-sided ("marriage") and general pairing ("roommate") problems. We introduce "sequences," lists of matchings that are repeated in order. Stable sequences are natural extensions of stable matchings; case in point, we show that a sequence of stable matchings is stable. In addition, stable sequences can provide solutions to problems for which stable matchings do not exist. In a sense, they allow us to "balance" the interest of the agents at different matchings. In this way, sequences can be superior to matchings in terms of welfare and fairness.

A seminal result due to Roth (1982, Math Oper Res 7(4), 617–628) is that no *strategy-proof* rule always selects stable matchings. In contrast, we show that there is a *weakly group sd–strategy-proof* rule that selects stable sequences. We call it the *Compromises and Rewards* rule, *CR*. We find that stronger incentive properties are incompatible with much weaker stability properties and vice versa. The *CR* rule satisfies two fairness axioms: *anonymity* and *side-neutrality*. For the general problem, the *Generalized CR* rule is *sd–5-stable* (cannot be blocked by groups of five or fewer agents), *weakly sd–strategy-proof*, and *anonymous*. In addition, the *Extended All-Proposing Deferred Acceptance* rule is *sd-stable*, *anonymous*, and *individually rational at all times* on a restricted domain. We provide a condition under which our results still hold if agents have cardinal preferences and compare sequences using "expected utility."

Keywords: Pairing problems, sequences, stability, strategy-proofness, algorithms
JEL: C62, D02, D60

[☆] I am especially thankful to Tommy Andersson, Albin Erlanson, and Vikram Manjunath for many helpful discussions. Comments by Lars Ehlers, Ehud Kalai, Thayer Morrill, Michael Ostrovsky, David C. Parkes, Alexander Reffgen, Alexander Teytelboym, William Thomson, M. Utku Ünver, Rajiv Vohra are gratefully appreciated. The paper has also benefited from comments made by participants at seminars at University of Rochester and Lund University, as well as participants at the Workshop on Coalitions and Networks (Montreal) and the meeting of Social Choice and Welfare (Boston). Part of the paper was written while I was visiting Boston College and University of Rochester. I thank the respective economics departments for their hospitality. Remaining errors are my own. Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged.

November 22, 2014

1. Introduction

Can we pair agents in a stable way *and* ensure that they never lie about their preferences? In a static one-shot setting, the answer is "no" (Roth, 1982). But what about a dynamic setting in which the agents interact repeatedly and can switch matchings from time to time? To answer this question, we introduce *sequences*, lists of matchings that are repeated over time. In sharp contrast to Roth (1982), we find *non-manipulable rules that select stable sequences*.

Our starting point is that sometimes the solution to a matching problem is not just one particular matching. Consider for instance the following challenge facing physicians worldwide. A doctor's emotional detachment from her patient has long been considered a necessity to prevent a loss of objectivity and perspective in his treatment (Blumgart, 1964). Case in point, the American College of Physicians (Snyder, 2012, page 81) makes the following recommendation:

> *Physicians should usually not enter into the dual relationship of physician-family member or physician-friend for a variety of reasons. The patient may be at risk of receiving inferior care from the physician. Problems may include effects on clinical objectivity, inadequate history taking or physical examination, overtesting, inappropriate prescribing, incomplete counseling on sensitive issues, or failure to keep appropriate medical records.*

Arguably, remaining detached gets more difficult over time as the patient and doctor get increasingly familiar. The two connect for instance through small talk during visits and, in some cases, interactions on social media websites (Bosslet et al., 2011). However, the entire issue is likely less severe if any subpar treatment is quickly detected. For this purpose, consider an arrangement in which a patient primarily sees a "main" doctor but occasionally gets a second opinion from some "reserves." In our context, this is modelled as a sequence consisting of an oft-occurring default matching that at times is swapped for "check-up" matchings. In this way, doctors monitor each other, making sure the treatment never deteriorates too far.

Imagine next being the owner of some shops that are operated by pairs of employees, say by a chef and a cashier. Surely, the employees need to be comfortable working alongside each other, but this can lead to a negative shirking effect. Essentially, if the perceived likelihood of one getting caught shirking is smaller, one rationally provides less effort (see for instance Shapiro and Stiglitz, 1984, page 439).[1] As the owner, you may therefore wish to construct a varied schedule for your employees to maintain a good rotation. The "stability" of this schedule depends crucially on how you weigh "compromises and rewards" for the agents. You cannot always pair an employee with someone she dislikes (have her "compromise") as she may then seek employment elsewhere. However, you may be forced to *occasionally* match her to someone she dislikes. If you do so, you should make sure she is "rewarded" for this by sometimes

[1] It is reasonable to believe that the better friends you are with your coworker, the lower the risk that she tells on you. In this way, friends can more easily collude to promote their own interest ahead of that of the firm (see also Tirole, 1986; Laffont and Tirole, 1993). We do acknowledge that friendship can have the opposite effect. If two individuals have to paint a fence, then one shirking implies more work for the other. In this case, friends may be less inclined to shirk. See also Mas and Moretti (2009).

matching her with someone she likes better. Keep in mind also that a reward for one agent may be a compromise for the partnering agent. Thus, designing a functional schedule is both challenging and important.

Finally, consider the employment of lifeguards at public swimming pools. There often needs to be more than one guard overseeing the pool at all times to ensure swimmers' safety.[2] Let us restrict attention to pools for which using more than two guards can be ruled out as unnecessary (alternatively, to pool complexes where pairs are responsible of keeping watch over smaller areas). The pairing frequently has to change to prevent the guards from becoming inattentive and "too comfortable" with one another:[3] The American Red Cross (2012, page 48) suggests that lifeguards should rotate stations every 20 to 30 minutes. When constructing the lifeguard schedule, you have to take into account the guards' preferences: where one sees a relaxed co-worker, another sees a lazy no-good slacker. If some guards keep getting paired unfavorably, they may decide to take their talents elsewhere. For instance, they may apply at a competitor, or even start their own lifeguard venture.

In this paper, we look for sequences from which no agents can benefit from deviating, that is, *stable* sequences. Besides being desirable solutions to many problems, sequences are interesting to study as they have several nice properties. For instance, a "well-balanced" sequence of matchings can Pareto-dominate a stable matching (Example 1). As just eluded to, sequences allow us to promote different agents at different matchings. Thus, they can provide a more fair pairing over time than any one particular matching (Example 2). Finally, an important difference to matchings is that there are non-manipulable rules that select stable sequences (Example 3).

We analyze two-sided and general pairing problems.[4] Hence, there is a set of agents and we wish to pair them. For the two-sided problem, the agents are divided into two groups and matchings are restricted to be of agents from different groups. Each agent has a preference over whom she is matched to and compares sequences using stochastic dominance (sd) comparisons. That is, an agent finds the sequence Σ at least as desirable as the sequence Ψ if she is matched at least as frequently with her most preferred agent in Σ than in Ψ; at least as frequently with her two most preferred agents in Σ than in Ψ; and so on. There is no money in the model: hence, one cannot trade-off a bad schedule for a pay rise.[5] The agents report their preferences to a central authority which selects a sequence (think of the employer proposing a schedule to her employees).[6] We design desirable rules for making this selection.

[2] See for instance South Carolina state law act number 159 of year 2012. Alternatively, the Connecticut Office of Legislative Research (2012) presents a survey on `http://www.cga.ct.gov/2012/rpt/2012-R-0524.htm`.

[3] Griffiths (2002) notes that "Monotony leads to boredom, which, in turn, leads to a lack of vigilance, one of the biggest problems in lifeguarding today."

[4] These problems are usually referred to as the "marriage" and the "roommate" problem. We find that the word "marriage" has connotations irrelevant for most real-world situations that can be modelled as "marriage" problems. It is certainly controversial to talk about "sequences" in marriage. Therefore, we choose a neutral name.

[5] The two-sided pairing problem with money is the "assignment game" of Shapley and Shubik (1971). The general version is the "partnership formation problem" studied by Talman and Yang (2011), Alkan and Tuncay (2014), and Andersson et al. (2014a,b) among others. Similar to this paper, using compromises to resolve instability in the partnership formation problem is done by Gudmundsson (2013).

[6] Stability is also essential in *decentralized* settings where agents interact repeatedly. Once a stable matching

The appeal of our rules is intimately connected to the axioms they satisfy. Therefore, perhaps in abundance, we spend some time on explaining the importance of stability and strategy-proofness. At the core of most successful applications of matching theory to real-world problems is the insight nicely summarized by Roth (2002).[7] Namely, rules that select stable matchings tend to stay in use year in, year out, whereas others do not. To see why, consider a procedure that yields an unstable matching. By definition, there are agents who can do better on their own, circumventing the procedure. Even if the procedure has nice properties in terms of, say, efficiency and fairness, these properties will immediately be harmed if some agents do not participate. Even more troubling is that once it becomes clear that one can benefit from bypassing the official system, others may follow suit – possibly leading to a breakdown of the entire procedure.[8,9]

Thus, a lack of stability can turn an otherwise well-designed rule useless in practice. The property's appeal can however be in question if it does not come bundled with the following characteristic. As the preferences are not available to us, the agents must themselves provide them. It is fundamental for a stable rule that the reported preferences are the true ones. Otherwise, the rule selects an outcome stable *with respect to the wrong preferences*, but perhaps *unstable* with respect to the true ones. Our rules should be *strategy-proof*: no agent should ever benefit from lying about her preference.[10]

Recall, Roth (1982) shows that *no strategy-proof rule always selects stable matchings*. Thus, the properties we just argued are essential are incompatible.[11] As a second-best solution, much research has since been focused on the *Deferred Acceptance* rule of Gale and Shapley (1962), *DA* for short. This rule's success is rooted in breaking the symmetry between the two sides of agents, making one side "propose" to the other. The rule is manipulable, but only by agents receiving proposals.[12] It selects stable matchings, but favors the proposers in this selection. This may at first sight seem unfair: the choice of proposing agents induces a potentially significant welfare gap between the two sides. However, remember that strategy-

is reached, myopic agents have no incentives to deviate from it. Therefore, it is interesting to know whether individual agents, acting in their own self-interest, can coordinate to reach a stable matching (Roth and Vande Vate, 1990; Abeledo and Rothblum, 1995; Diamantoudi et al., 2004; Inarra et al., 2008). For future research, we may ask whether agents that make plans over a longer time horizon can coordinate to reach a stable sequence.

[7] See the literature on kidney exchange (Roth et al., 2004), school choice (Abdulkadiroğlu and Sönmez, 2003), and the National Resident Matching Program (Roth and Peranson, 1999). Some surveys of the literature are Roth and Sotomayor (1990), Sönmez and Ünver (2010), and Manlove (2013).

[8] This "unraveling" in matching markets is examined by Niederle and Roth (2003) and Ostrovsky and Schwarz (2010) among others.

[9] Stability can also be viewed as a fairness property. What a group of agents can achieve on their own ought to be a lower bound on what they are assigned. An outcome is stable precisely when all lower bounds are met.

[10] Strategy-proofness can also be viewed as a fairness property. As some agents may be more strategic than others, strategy-proofness can "level the playing field" for sincere (non-strategic) and sophisticated (strategic) agents (Pathak and Sönmez, 2008).

[11] Positive results have been found on restricted preference domains as these limit the possibilities for manipulation, see for instance Alcalde and Barberà (1994) and Akahoshi (2014).

[12] A strengthening of Roth's (1982) result is that strategy-proof rules either are inefficient or not individually rational (Alcalde and Barberà, 1994). The *DA* rule is in a sense a compromise. It allows for a limited amount of manipulation in exchange for efficiency and some fairness.

proofness is *a necessity* for any type of fairness. The outcome selected by a manipulable rule designed to be fair may, if agents frequently misreport their preferences, not be very fair. Therefore, we have had to rely on rules that prioritize incentive constraints over fairness in this way.

There are similar incompatibilities for rules that select sequences. If we insist that the rule always selects a sequence of stable matchings, then we have to give up on the very weakest incentive properties. Taking the opposite role, if we require an *sd–strategy-proof* rule, the we have to give up on the very weakest stability properties. This is summarized in Theorem 1. However, our results also indicate that sequences provide a different point of view. In Theorem 2, we present an *sd-stable*, *weakly group sd–strategy-proof*, *anonymous*, and *side-neutral* rule for selecting sequences for the two-sided problem.[13] Hence, not only can we combine stability with strategy-proofness, we can strengthen the latter to protect against collusive behaviour *and* we can do this without treating the agents on the two sides asymmetrically. The rule we develop is the *Compromises and Rewards* rule, *CR* for short. It is based in a novel way on David Gale's *Top Trading Cycles* algorithm (Shapley and Scarf, 1974). We also find that a *weakly sd–strategy-proof* rule occasionally selects a sequence that contains unstable matchings – even though there always are stable matchings. For the general problem, we first illustrate how stable sequences are natural extensions of stable matchings. More precisely, in Theorem 3 we establish that a sequence of stable matchings is stable. This can be especially useful if we wish to achieve a "fair" pairing as the welfare of the agents may differ significantly at different stable matchings. We then construct an intuitive extension of the *DA* rule. For the original *DA* rule, the two sides of agents are essential. Our extension is handcrafted for sequences and does not require the agents to be divided. In Theorem 4, we show that the rule is *sd-stable*, *anonymous*, and *individually rational at all times* on a restricted domain of problems. The domain strictly subsumes the domain of problems that have stable matchings. In Theorem 5, we show that the *Generalized CR* rule is *sd–5-stable*,[14] *weakly sd–strategy-proof*, and *anonymous*. We conjecture that it is (fully) *sd-stable* and *weakly group sd–strategy-proof*.

In Section 7, we describe how all results can be generalized to a larger domain of preferences. More precisely, we provide a condition under which our results still hold if agents have cardinal preferences and compare sequences using "expected utility." The condition has a natural interpretation and the preferences are complete (in contrast to the stochastic dominance relation). In relation to the existing literature, this is the first paper to examine strategy-proofness specifically for the general pairing problem.[15]

Results of a similar nature to Theorem 3 have been found in other papers; there are many solution concepts that extend stable matchings. Ours generally does not pinpoint one matching, but rather a group of them. Different matchings need not occur with the same frequency in a sequence. In this way, stable sequences give a more precise prediction than a set-valued

[13] All properties are defined formally in Subsection 2.3.

[14] A rule is *sd–k-stable* if no group of at most k agents can block its suggested sequence.

[15] Strategy-proof "single-lapping rules" have been studied for coalition formation problems on restricted domains, see Pápai (2004) and Rodríguez-Álvarez (2009). These problems are more general than pairing problems.

solution concept like von Neumann-Morgenstern stable sets (von Neumann and Morgenstern, 1947; Klaus et al., 2011) or absorbing sets (Inarra et al., 2010) (for these, all matchings are attributed equal importance). However, they are not as "exact" as a farsightedly stable or a stochastically stable matching (see e.g Klaus et al., 2010). It does retain a sense of "full stability", in contrast to "almost stable" (Abraham et al., 2006) and "maximum stable" matchings (Tan, 1990) which rather are focused on finding the least unstable matchings.[16]

We mention also that we can interpret a sequence as a lottery over matchings. This opens up the scope of applications even further. For instance, to decide on which students to accept to a course, a lottery can be used to break potential ties.[17] A similar type of random tie breaking is used in some kidney exchange mechanisms (Ashlagi et al., 2013).

The paper is structured as follows. In Section 2, we present the model. In Section 3, we provide some motivating examples. In Section 4, we present results for two-sided problems. In Section 5, we examine many-to-one problems. In Section 6, we present results for general problem. In Section 7, we discuss fractional and probabilistic matchings as well as a generalization of the preference domain. In Section 8, we conclude. Proofs, auxiliary results, and additional details are postponed to the Appendix.

2. Model and definitions

2.1. Preliminaries

There are n **agents** N divided into sets M and W of equal size, $N = M \cup W$.[18] A **matching** is $\mu\colon N \to N$ such that $\mu(i) = j \Leftrightarrow \mu(j) = i$ for each $\{i, j\} \subseteq N$. Matchings are restricted to be between agents of different groups. That is, for each $m \in M$, $\mu(m) \in W \cup \{m\}$, and for each $w \in W$, $\mu(w) \in M \cup \{w\}$. If $\mu(i) = i$ for some $i \in N$, then i is **single** at μ. With some abuse of notation, for each $S \subseteq N$, $\mu(S)$ denotes the set of partners of agents in S at μ: $\mu(S) = \{\mu(i) : i \in S\}$. We often describe a matching by its graph, that is, $\mu = \{(i, j), (k, m), \ldots\}$ is short for $\mu(i) = j$, $\mu(k) = m$, and so on. The **set of matchings** is $\mathscr{M}$. A **preference** for $i \in N$ is the binary relation R_i on N such that i finds j at least as desirable as k whenever $j\ R_i\ k$. Preferences are strict; formally, if $j\ R_i\ k$ and $k\ R_i\ j$, then $j = k$. The strict relation is denoted P_i. We do not rule out that $i\ P_i\ j$, that is, matching with some agent j may be less desirable to i than being single. In addition, M-agents prefer W-agents to other M-agents and vice versa. Formally, for each $m \in M$, $m' \in M \setminus \{m\}$, $w \in W$, and $w' \in W \setminus \{w\}$, $w\ P_m\ m'$ and $m\ P_w\ w'$. The **set of preferences** is denoted $\mathscr{R}$.[19] A **profile of preferences** is $R \equiv (R_i)_{i \in N} \in \mathscr{R}^n$. A two-sided pairing problem, or simply a **two-sided problem**, is completely described by $R \in \mathscr{R}^n$. A group of agents **block** a matching if they can pair up in a way that makes everyone in the group at least as well off

[16] General pairing problems that allow stable matchings are characterized by Tan (1991). The problem with weak preferences (allowing for indifferences) is considered by Gudmundsson (2014).

[17] This is for instance the case in the Swedish admission system, see `https://www.antagning.se/sv/Ta-reda-pa-mer-/Platsfordelning-och-urval/Vid-lika-meritvarde/` (in Swedish).

[18] This is without loss of generality. If M and W are not of the same size, we can extend the smaller of them with "null agents" who prefer being single. This only comes into play when defining *side-neutrality*.

[19] Formally, we impose three additional constraints on the preference R_i. R_i is *reflexive*: $i\ R_i\ i$; *complete*: for each $j, k \in N$, $j\ R_i\ k$ or $k\ R_i\ j$; and *transitive*: $j\ R_i\ k$ and $k\ R_i\ m$ implies $j\ R_i\ m$.

and someone better off. Thus, $S \subseteq N$ blocks $\mu \in \mathscr{M}$ if there is $\mu' \in \mathscr{M}$ such that $\mu'(S) = S$, for each $i \in S$, $\mu'(i)\ R_i\ \mu(i)$, and for some $j \in S$, $\mu'(j)\ P_j\ \mu(j)$. A matching is **stable** if no group blocks it. For $R \in \mathscr{R}^n$, the **set of stable matchings** is $\mathscr{C}(R)$. The set $\mathscr{C}(R)$ has a lattice structure with respect to R with an M-optimal matching μ_R^M and a W-optimal matching μ_R^W (Gale and Shapley, 1962). That is, for each $\mu \in \mathscr{C}(R)$, each $m \in M$, and each $w \in W$, $\mu_R^M(m)\ R_m\ \mu(m)$ and $\mu_R^W(w)\ R_w\ \mu(w)$. A matching is **Pareto-efficient** if no other matching makes everyone at least as well off and someone better off. Equivalently, $\mu \in \mathscr{M}$ is Pareto-efficient if N does not block μ.

To formalize our algorithms we require some additional definitions. First, the injective function $r\colon N \to \mathbb{R}$ associates to each $i \in N$ her **rank** $r(i)$. Throughout, r is fixed. Second, take as given a directed graph (V, E), where V is its vertex set and E its edge set. A **cycle** in (V, E) is a list of vertices such that each vertex has an outgoing edge to the succeeding vertex. The cycle is completed by having the last vertex pointing to the first. For technical reasons, we describe a cycle by putting its top-ranked agent first. Thus, a cycle of length m in (V, E) is $C = (c_1, c_2, \ldots, c_m)$ such that, for each $c_k \in C$, $c_k \in V$, $(c_k, c_{k+1}) \in E$ (mod m), and $r(c_1) \geq r(c_k)$. A **loop** is a cycle of length 1.

2.2. Sequences

A sequence of matchings, a **sequence** for short, is a finite list $[\mu_1, \mu_2, \ldots]$ such that $\mu_t \in \mathscr{M}$ for each t.[20] Denote the **set of sequences** $\mathscr{S}$. Each $\Sigma \in \mathscr{S}$ induces an $n \times n$ matrix σ, where, for each $\{i, j\} \subseteq N$, $\sigma_{ij} \in [0, 1]$ is the frequency at which i is matched to j in Σ.

$$\sigma_{ij} = \frac{\#\{\mu \in \Sigma : \mu(i) = j\}}{\#\{\mu \in \Sigma\}}$$

Sequences that induce the same matrices are **equivalent** and viewed as equal by the agents. This rules out time discounting: the sequences $[\mu_1, \mu_2]$, $[\mu_1, \mu_1, \mu_2, \mu_2]$, and $[\mu_1, \mu_2, \mu_1, \mu_2]$ are for instance equivalent. In addition, there are no matching externalities.[21] That agents do not discount over time is a sensible approximation if the agents are patient and the sequences are relatively short in relation to the full time horizon (which here is infinite as the sequences are repeated indefinitely). We denote by $\#\Sigma$ the length of the shortest sequence equivalent to Σ.[22]

Agent i's **preference over sequences** is the binary relation R_i^{sd} on $\mathscr{S}$ induced by R_i. For each $\Sigma \in \mathscr{S}$ associated with σ and each $\Psi \in \mathscr{S}$ associated with ψ,

$$\Sigma\ R_i^{\text{sd}}\ \Psi \Leftrightarrow \forall k \in N,\ \sum_{jR_ik} \sigma_{ij} \geq \sum_{jR_ik} \psi_{ij}.$$[23]

[20] The assumption of sequences being finite is not a necessary one. It is made to indicate that we consider a sequence to be something that is repeated.

[21] Consider the following matchings: $\mu_1 = \{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$, $\mu_2 = \{(m_1, w_2), (m_2, w_3), (m_3, w_1)\}$, $\mu_3 = \{(m_1, w_3), (m_2, w_1), (m_3, w_2)\}$, $\mu_1' = \{(m_1, w_1), (m_2, w_3), (m_3, w_2)\}$, $\mu_2' = \{(m_1, w_2), (m_2, w_1), (m_3, w_3)\}$, $\mu_3' = \{(m_1, w_3), (m_2, w_2), (m_3, w_1)\}$. Then the sequences $[\mu_1, \mu_2, \mu_3]$ and $[\mu_1', \mu_2', \mu_3']$ are equivalent, even though they share no matching. That agents care only about their partners and not the pairing of the others is a standard assumption. Some exceptions are Sasaki and Toda (1996) and Gudmundsson and Habis (2013).

[22] The # symbol is also used to indicate set cardinalities and cycle lengths.

[23] The notation $\sum_{jR_ik}$ is short for summing over the set $\{j \in N : j\ R_i\ k\}$.

In words, an agent finds $\Sigma \in \mathscr{S}$ at least as desirable as $\Psi \in \mathscr{S}$ if she is matched at least as frequently with her most preferred agent in Σ than in Ψ; at least as frequently with her two most preferred agents in Σ than in Ψ; and so on. The strict relation is P_i^{sd} and the indifference relation is I_i^{sd}. Note that R_i^{sd} is an incomplete relation: there are sequences the agent cannot compare. We do not exploit this peculiarity of the preferences to prove any of our positive results. In fact, we describe how to strengthen our results by "completing the preferences" in Section 7.

The group $S \subseteq N$ sd-blocks, or simply **blocks**,[24] $\Sigma \in \mathscr{S}$ if there is $\Psi \in \mathscr{S}$ such that $\mu'(S) = S$ for each $\mu' \in \Psi$, for each $i \in S$, $\Psi\ R_i^{\text{sd}}\ \Sigma$, and for some $j \in S$, $\Psi\ P_j^{\text{sd}}\ \Sigma$. A sequence is sd-stable, **stable** for short, if no group blocks it. For $R \in \mathscr{R}^n$, the **set of stable sequences** is $\mathscr{C}^{\text{seq}}(R)$. For $k \leq n$, $\Sigma \in \mathscr{S}$ is sd–k-stable, or simply ***k*-stable**, if no $S \subseteq N$ such that $\#S \leq k$ blocks Σ. In settings where it is difficult for large groups to coordinate, k-stable sequences for $k < n$ are interesting to study. The **set of *k*-stable sequences** is $\mathscr{C}_k^{\text{seq}}(R)$. A sequence is **sd-efficient** if no other sequence makes everyone at least as well off and someone better off.[25] Equivalently, $\Sigma \in \mathscr{S}$ is sd-efficient if N does not block Σ. If $\Sigma \in \mathscr{S}$ is stable, then Σ is k-stable for all k and Σ is sd-efficient. If $\Sigma \in \mathscr{S}$ is sd-efficient, then each $\mu \in \Sigma$ is Pareto-efficient.

For a group to block, at least one agent in the group has to be made better off. This is a standard assumption. For the remaining agents in the group, we require that all should be at least as well off. If this is weakened, in the sense that we only require that no agent should be worse off, blocking is made easier and the notion of stability is made stronger. Formally, $S \subseteq N$ **weakly sd-blocks** $\Sigma \in \mathscr{S}$ if there is $\Psi \in \mathscr{S}$ such that $\mu'(S) = S$ for each $\mu' \in \Psi$, for each $i \in S$, $\Sigma\ P_i^{\text{sd}}\ \Psi$ is not true, and for some $j \in S$, $\Psi\ P_j^{\text{sd}}\ \Sigma$. A sequence is **strongly sd-stable** if no group weakly sd-blocks it. Example 7 can be used to show that some general pairing problems do not have strongly sd-stable sequences (though they always exist for two-sided problems).

There is a non-cooperative foundation to the idea that agents block the entire sequence rather than individual matchings in it.[26] Consider the sequence as an agreement between the agents. Deviation away from this agreement is punished by the other agents, in the sense that a deviating group S no longer gets to match with anyone in $N \setminus S$. Then S would deviate if and only if they can block as above. The underlying idea here therefore bears similarities to folk theorems in game theory that state that cooperation can be sustained in repeated interactions through threats of long-term punishments to deviating agents.

For pairing problems, a matching is stable whenever there is no pair of agents that can block it (that is, whenever it is 2-stable). This makes it straightforward to check the stability of a matching. Here, a sequence is 2-stable whenever no group can block it to a sequence consisting of just one matching (a formal proof of this claim for the general problem is provided in Appendix B, Proposition 1), though 2-stable sequences need not be stable. Hence, it is a more complicated task to determine whether a sequence is stable or not.[27]

[24] The overlapping terminology should not cause any confusion. We always specify whether it is a matching or a sequence that is blocked. Similarly, stable sequences are easy to distinguish from stable matchings.

[25] Bogomolnaia and Moulin (2001) call this "ordinally efficient". We follow Thomson's (2013) recommendation.

[26] If we allow agents to block individual matchings without affecting the remainder of the sequence, the set of stable sequences coincides with those that only contain stable matchings (see also Theorem 3).

[27] This is not saying that it is "difficult" for $S \subseteq N$ to know whether they block a sequence. For each of the $\#S$ agents

2.3. Rules and desirable properties

A **rule** is a mapping $\varphi\colon \mathscr{R}^n \to \mathscr{S}$. A rule φ is **manipulable** at $R \in \mathscr{R}^n$ by $i \in N$ if there is a lie $R'_i \in \mathscr{R}$ such that $\varphi(R'_i, R_{-i})\ P_i^{\text{sd}}\ \varphi(R)$.[28] A rule φ is **weakly sd–strategy-proof** if, for each profile, each agent is not better off lying than she is reporting her true preference,

$$\forall R \in \mathscr{R}^n,\ \forall i \in N,\ \forall R'_i \in \mathscr{R},\ \varphi(R'_i, R_{-i})\ P_i^{\text{sd}}\ \varphi(R) \text{ is not true.}$$

A rule φ is **sd–strategy-proof** if, for each profile, each agent is at least as well off reporting her true preference as she is lying,

$$\forall R \in \mathscr{R}^n,\ \forall i \in N,\ \forall R'_i \in \mathscr{R},\ \varphi(R)\ R_i^{\text{sd}}\ \varphi(R'_i, R_{-i}).$$

This is a strengthening as the relation R_i^{sd} is incomplete. A rule φ is **manipulable to the top** at $R \in \mathscr{R}^n$ by $i \in N$ if, when telling the truth, i is not always matched with her most preferred agent j, whereas she is when telling a lie $R'_i \in \mathscr{R}$:

$$\exists \mu \in \varphi(R) \text{ s.t. } \mu(i) \neq j \text{ and } \forall \mu' \in \varphi(R'_i, R_{-i}),\ \mu'(i) = j, \text{ where, } \forall k \in N,\ j\ R_i\ k.$$

A rule φ is **group-manipulable** at $R \in \mathscr{R}^n$ by $S \subseteq N$ if there is $R'_S \equiv (R'_i)_{i \in S} \in \mathscr{R}^S$ such that, for each $i \in S$, $\varphi(R'_S, R_{-S})\ R_i^{\text{sd}}\ \varphi(R)$, and for some $j \in S$, $\varphi(R'_S, R_{-S})\ P_j^{\text{sd}}\ \varphi(R)$. A rule φ is **weakly group sd–strategy-proof** if φ never is group-manipulable. The incentive properties are logically related as follows.

$$\begin{aligned} &\varphi \text{ is } \textit{sd–strategy-proof} \text{ or } \textit{weakly group sd–strategy-proof} \\ &\quad \Rightarrow \varphi \text{ is } \textit{weakly sd–strategy-proof} \\ &\qquad \Rightarrow \varphi \text{ is } \textit{not manipulable to the top.} \end{aligned}$$

A rule φ is **sd-stable** if it always selects a stable sequence:

$$\forall R \in \mathscr{R}^n,\ \varphi(R) \in \mathscr{C}^{\text{seq}}(R).$$

A rule φ is **stable at all times** if it always selects a sequence of stable matchings:

$$\forall R \in \mathscr{R}^n,\ \forall \mu \in \varphi(R),\ \mu \in \mathscr{C}(R).$$

If φ is *stable at all times*, then φ is *sd-stable* (Manjunath, 2013, Proposition 3). A rule φ **respects mutual best** if it always matches agents who prefer one another to everyone else:

$$\forall R \in \mathscr{R}^n,\ \forall \{i, j\} \subseteq N \text{ s.t. } \forall k \in N,\ j\ R_i\ k \text{ and } i\ R_j\ k \Rightarrow \forall \mu \in \varphi(R),\ \mu(i) = j.$$

---

there are $\#S$ linear inequalities that need to be satisfied. By introducing slack variables, this boils down to solving a linear program where the number of inequalities grows quadratically in the number of agents. Whether a sequence is stable or not is computationally more difficult as the number of groups S grows exponentially.

[28] Except for i, agents have the same preferences at R as at (R'_i, R_{-i}). Agent i has changed her preference to R'_i. For $S \subseteq N$, (R'_S, R_{-S}) denotes a similar change of preference but for all agents in S.

For $k \leq n$, φ is **sd–k-stable** if it always selects a k-stable sequence: for each $R \in \mathscr{R}^n$, $\varphi(R) \in \mathscr{C}_k^{\text{seq}}$. In the literature, a 1-*stable* rule is usually referred to as "individually rational." We can strengthen this as follows. A rule φ is **individually rational at all times** if it always matches agents that find each other at least as desirable as being single:

$$\forall R \in \mathscr{R}^n,\ \forall i \in N,\ \forall \mu \in \varphi(R),\ \mu(i)\ R_i\ i.$$

If φ is *stable at all times*, then φ is *individually rational at all times*. The stability properties are logically related as follows.

$$\begin{gathered}\varphi \text{ is } \textit{stable at all times} \Rightarrow \varphi \text{ is } \textit{sd-stable} \Rightarrow \varphi \text{ is } \textit{sd–}k\textit{-stable} \text{ for } k \geq 2\\ \Rightarrow \varphi \text{ is } \textit{sd–}k'\textit{-stable} \text{ for } k' \leq k,\ k' \geq 2 \Rightarrow \varphi \textit{ respects mutual best.}\end{gathered}$$

Let Π be the set of all bijections on N. We use $\pi \in \Pi$ to permute the names of the agents, possibly changing the agents' membership from M to W or vice versa. Let $\Pi^S \subset \Pi$ be all "side-swapping" permutations. Formally, $\pi \in \Pi^S$ if and only if $i \in M \Leftrightarrow \pi(i) \in W$. Let $\Pi^P \subset \Pi$ be all "side-preserving" permutations. Formally, $\pi \in \Pi^P$ if and only if $i \in M \Leftrightarrow \pi(i) \in M$.

Consider $\pi \in \Pi$ applied to $R \in \mathscr{R}^n$. This yields a new profile $\tilde{R} \equiv \pi \circ R$ defined as follows. For all $\{i, j, k\} \subseteq N$, $j\ R_i\ k \Leftrightarrow \tilde{j}\ \tilde{R}_{\tilde{i}}\ \tilde{k}$, where $\tilde{i} \equiv \pi(i)$, $\tilde{j} \equiv \pi(j)$, and $\tilde{k} \equiv \pi(k)$. Applying $\pi \in \Pi^S \cup \Pi^P$ to a matching yields a different matching. For $\mu \in \mathscr{M}$, $\pi \circ \mu$ is defined as follows. A typical agent $i \in N$ becomes $\pi(i)$; her partner at μ is $\pi(\mu(i))$. Therefore $(\pi \circ \mu)(\pi(i)) = \pi(\mu(i))$. For $\Sigma \in \mathscr{S}$, $\pi \circ \Sigma = [\pi \circ \mu : \mu \in \Sigma]$ is the sequence obtained when applying $\pi \in \Pi^S \cup \Pi^P$ to each matching in Σ. A rule is **side-neutral** (in welfare terms) if neither side receives any "special treatment":

$$\forall R \in \mathscr{R}^n,\ \forall i \in N,\ \forall \pi \in \Pi^S,\ \varphi(\pi \circ R)\ I_i^{\text{sd}}\ \pi \circ \varphi(R).$$

A rule is **anonymous** (in welfare terms) if no agent receives any "special treatment":

$$\forall R \in \mathscr{R}^n,\ \forall i \in N,\ \forall \pi \in \Pi^P,\ \varphi(\pi \circ R)\ I_i^{\text{sd}}\ \pi \circ \varphi(R).$$

3. Three motivating examples

We next present some examples that highlight potential advantages of sequences over matchings. In the first, we use a sequence to Pareto-improve a stable matching.[29]

Example 1: Pareto-improving matchings through a sequence. Consider a scenario in which each agent cares much more about their most preferred partner than they care about the others. Intuitively, an agent may then be willing to exchange some time with a "middle-ranked" partner for more time with a lower ranked partner *and* more time with their most preferred partner. The main point of this example is *not* to convince you that these are generic and natural preferences. Rather it is to highlight that in the event that agents happen to have these tastes – a possibility that cannot a priori be ruled out – sequences add options not present if we focus only on single matchings. For this example, the agents are $M = \{m_1, m_2\}$ and $W = \{w_1, w_2\}$ with preferences in Table 1.

[29] The example can also be found in Manjunath (2013).

R_{m_1}	R_{m_2}	R_{w_1}	R_{w_2}	u
w_1	w_2	m_2	m_1	4
$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	2
w_2	w_1	m_1	m_2	1

Table 1: Preferences for Example 1. The $\emptyset$ symbol represents being single. The column u displays (cardinal) utility levels compatible with the story presented.

At the unique stable matching μ every agent is single, $\mu(i) = i$ for all $i \in N$. The sequence $[\mu]$ – that is, spending all time matched according to μ – is stable. However, there are additional stable sequences. Indeed, this holds with few exceptions: there are generally many more stable sequences than there are stable matchings. For the sake of argument, consider the sequence $[\mu_1, \mu_2]$, where $\mu_1 = \{(m_1, w_1), (m_2, w_2)\}$ and $\mu_2 = \{(m_1, w_2), (m_2, w_1)\}$. To be sure, neither μ_1 nor μ_2 is stable – so why don't w_1 or w_2 block μ_1? Recall, once a sequence is interrupted, it has to be replaced by a new sequence. Agents cannot change a matching in the sequence and expect the future plans to remain intact. In this case, if w_1 were to block μ_1, then μ_2 may not be formed tomorrow. This would be upsetting for w_1. Keep in mind also that sequences are repeated indefinitely. Agents m_1 and m_2 have no desire to block μ_2 as the perceived sequence at μ_2 is $[\mu_2, \mu_1]$. Strikingly, those agents who can block μ_1 are matched to their first choice at μ_2 and vice versa. As each agent finds their most preferred agent much better than the others, the agent prefers this alternating "compromise" and "reward" to μ. ∘

Next, we show a fairness issue that arises when focusing only on single matchings. Namely, stable matchings can favor some agents at the expense of others. To be sure, a matching is not stable because everyone is happy – it is stable because those unsatisfied cannot convince their preferred agents to match with them rather than with their current partners.

Example 2: Unbalancedness of stable matchings. Consider the simplest of examples that showcases the opposing interests of M and W. The two-sided problem consists of agents $M = \{m_1, m_2\}$ and $W = \{w_1, w_2\}$ with preferences in Table 2.

R_{m_1}	R_{m_2}	R_{w_1}	R_{w_2}
w_1	w_2	m_2	m_1
w_2	w_1	m_1	m_2
$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$

Table 2: Preferences for Example 2.

There are two stable matchings: $\mu_1 = \{(m_1, w_1), (m_2, w_2)\}$ and $\mu_2 = \{(m_1, w_2), (m_2, w_1)\}$. The agents in M are well off at μ_1, matching to their most preferred agents. The agents in W are worse off, matching to their least preferred agents. The opposite holds for μ_2. As μ_1 and μ_2 are the only Pareto-efficient matchings, the set of stable sequences contains exactly

those that include no matching other than μ_1 and μ_2. For instance, it includes $[\mu_1, \mu_2]$, which arguably is a fair compromise for the agents. Thus, sequences can eliminate welfare gaps that exist only because we restrict ourselves to selecting one particular matching. ∘

As discussed in the introduction, there is no *strategy-proof* rule that always selects stable matchings. As it happens, this is equivalent to that no *strategy-proof* rule always selects 2-stable matchings. In contrast, in the following example we construct a rule that is *weakly group sd–strategy-proof* and *sd–3-stable*.[30]

Example 3: Some strategy-proofness and stability. Let φ be the rule that, to each $R \in \mathscr{R}^n$, selects the sequence obtained as follows. First, if possible, find a pair $\{i, j\} \subseteq N$ who prefer one another to all other agents. Formally, $j \; R_i \; k$ and $i \; R_j \; k$ for all $k \in N$. (Here, by "pair", we also consider *one* agent who prefers being single.) We match i and j at each matching in the sequence. As these agents cannot be made better off in any way, they will not be part of a manipulating or blocking group. Remove the pair and reiterate. We may now find a new pair $\{k, m\}$ who prefer each other to all agents in $N \setminus \{i, j\}$. We match k and m at each matching in the sequence. The only way for k and m to potentially improve is to match with the agents already dealt with (i and j), but, as we already established that these have no interest in changing partners, k and m cannot be made better off. Repeat until no more such pairs are found.[31]

Label the set of remaining agents S and order S arbitrarily. We create one matching "for" each agent in S. In particular, at the kth matching, match the kth agent of S to her most preferred agent in S. Leave everyone else in S single at the matching.

One easily finds that no agent in S can benefit from misreporting her preferences. Hence, the rule φ is *weakly sd–strategy-proof*. That no group of agents jointly can manipulate the rule requires a more involved argument. We refer to Figure 1 for a sketch of the proof for a particular two-sided problem. Lastly, it is immediate that S generally will block the sequence. To be more precise, any group of agents that form a cycle as described in Figure 1 can block. As such cycles must contain at least four agents, the rule is *sd–3-stable*.[32] ∘

4. Results on two-sided problems

Though the rule designed in Example 3 has several nice properties, it is inefficient and certainly not *sd-stable*. In this section, we first propose an *sd-stable* rule by using the *DA* mechanism. The rule is *stable at all times*, a property Example 4 shows implies that the rule is *manipulable to the top*. The example also show that no *sd–strategy-proof* rule *respects mutual best*. Thereafter, we present an *sd-stable* and *weakly group sd–strategy-proof* rule.

[30] The rule is also *side-neutral* and *anonymous*, but not *individually rational at all times*.

[31] For problems that are *α-reducible* (Alcalde, 1995), we can find such pairs in any subset of the agents. On the restricted domain of α-reducible problems, φ is *sd-stable*, even for general pairing problems.

[32] If there is a loop or a cycle of two agents, then that should have been processed in the previous step. Cycles are of even length as they alternate between M- and W-agents. Hence, the shortest cycle is of length 4 or more.

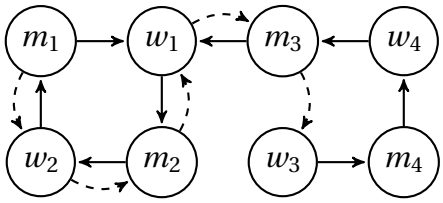

Figure 1: Let $S = \{m_1, w_1, m_2, w_2, \ldots, m_4, w_4\}$. A solid line from agent i to j indicates that i prefers j to all other agents in S. In general, the agents partition into multiple components. Each component contains one cycle, that is, a list of agents such that each points to the agent following her in the list, the last agent pointing to the first. Here, there is a unique cycle (m_1, w_1, m_2, w_2).

Suppose a group of agents can manipulate the rule. In particular, say m_3 reports a preference where w_3 is preferred. This is indicated by the dashed arrow from m_3 to w_3. For this misreport to be beneficial for m_3, m_3 still needs to be matched some time with w_1. Then w_1 needs to report a preference where m_3 is preferred. Similarly, m_2 needs to point to (prefer) w_1, w_2 to m_2, m_1 to w_2. For this to be in the interest of m_1, w_1 has to report a preference where she prefers m_1. However, w_1 has to point to m_3. Hence, m_1 will not benefit from the misreports, and therefore the group cannot manipulate the rule in this way.

4.1. The Repeated Deferred Acceptance with Alternating Proposers rule

Recall that the *M-proposing DA* mechanism generates μ_R^M, the M-optimal stable matching. The rule that, to each $R \in \mathscr{R}^n$, selects $\left[\mu_R^M\right]$ is *stable at all times*, and hence *sd-stable* (Manjunath, 2013, Proposition 3). It is also *individually rational at all times* and *anonymous*, but not *side-neutral*. The last property can be added by modifying the rule as follows.

Repeated *DA* with Alternating Proposers. For each $R \in \mathscr{R}^n$, $RDAAP(R) = \left[\mu_R^M, \mu_R^W\right]$.

This is an improvement in terms of efficiency and fairness, though not in terms of incentives. Namely, the sequence selected by *RDAAP* is stable with respect to the reported preferences. However, if agents misreport their preferences, the rule may select a sequence that is unstable with respect to the true preferences. If a rule is not strategy-proof, some agents may attempt to manipulate it even if they do not have enough information about the others' preferences to surely benefit from their misreport.[33] Indeed, agents occasionally can manipulate *RDAAP*. The following example, familiar from Roth (1982), highlights this as well as provides two interesting impossibilities. First, if we insist on selecting sequences on stable matchings, then the rule will be *manipulable to the top*. That is, it will not satisfy even the very weakest of our incentive properties. Second, if we insist on *sd–strategy-proofness*, then the rule will not *respect mutual best*. That is, it will not satisfy even the very weakest of our stability properties.

Example 4: Two important impossibilities. Consider the two-sided problem with agents $M = \{m_1, m_2, m_3\}$ and $W = \{w_1, w_2, w_3\}$ with preferences in Table 3.

Part I: Let the rule φ be *stable at all times*. With $\mu_1 = \{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$ and $\mu_2 = \{(m_1, w_2), (m_2, w_1), (m_3, w_3)\}$, we have $\mathscr{C}(R) = \{\mu_1, \mu_2\}$, $\mathscr{C}(R'_{m_1}, R_{-m_1}) = \{\mu_1\}$,

[33] For empirical evidence of this claim and experimental results, see Braun et al. (2010) and Pais et al. (2011).

R_{m_1}	R_{m_2}	R_{m_3}	R_{w_1}	R_{w_2}	R_{w_3}
w_1	w_2	w_1	m_2	m_1	m_1
w_2	w_1	w_2	m_1	m_2	m_2
w_3	w_3	w_3	m_3	m_3	m_3

R'_{m_1}	R'_{w_1}
w_1	m_2
w_3	m_3
w_2	m_1

R''_{m_1}	R''_{m_3}	R''_{w_2}	R''_{w_3}
w_3	w_3	m_3	m_3
w_1	w_1	m_1	m_1
w_2	w_2	m_2	m_2

Table 3: Preferences for Example 4.

and $\mathscr{C}(R'_{w_1}, R_{-w_1}) = \{\mu_2\}$. Therefore $\varphi(R'_{m_1}, R_{-m_1}) = [\mu_1]$, $\varphi(R'_{w_1}, R_{-w_1}) = [\mu_2]$, and, for each $\mu \in \varphi(R)$, $\mu \in \mathscr{C}(R) = \{\mu_1, \mu_2\}$. If $\mu_1 \in \varphi(R)$, then w_1 manipulates φ to the top at R by reporting R'_{w_1}. If $\mu_2 \in \varphi(R)$, then m_1 manipulates φ to the top at R by reporting R'_{m_1}. Hence, φ is *manipulable to the top*.

Part II: Let φ *respect mutual best* and, to obtain a contradiction, *sd–strategy-proofness*. Suppose there is $\mu \in \varphi(R'_{m_1}, R_{-m_1})$ such that $\mu(m_1) = w_2$. If m_1 reports R''_{m_1}, as φ *respects mutual best*, for each $\mu \in \varphi(R''_{m_1}, R_{-m_1})$, $\mu(m_1) = w_3$. Then $\varphi(R'_{m_1}, R_{-m_1})\ R'_{m_1}\ \varphi(R''_{m_1}, R_{-m_1})$ is not true. This contradicts φ being *sd–strategy-proof*. Therefore, there is no $\mu \in \varphi(R'_{m_1}, R_{-m_1})$ such that $\mu(m_1) = w_2$.

Suppose next there is $\mu \in \varphi(R'_{m_1}, R_{-m_1})$ such that $\mu(w_2) \neq m_2$. If w_2 reports R''_{w_2}, as φ *respects mutual best*, for each $\mu \in \varphi(R'_{m_1}, R''_{w_2}, R_{-\{m_1,w_2\}})$, $\mu(w_2) = m_2$. Thus, w_2 is better off by telling a lie and φ is not *sd–strategy-proof*. Therefore, for all $\mu \in \varphi(R'_{m_1}, R_{-m_1})$, $\mu(w_2) = m_2$.

Suppose next there is $\mu \in \varphi(R'_{m_1}, R_{-m_1})$ such that $\mu(w_1) \neq m_1$. If w_1 reports R'_{w_1}, as φ *respects mutual best*, for each $\mu \in \varphi(R'_{m_1}, R'_{w_1}, R_{-\{m_1,w_1\}})$, $\mu(w_1) = m_1$. Thus, w_1 is better off by telling a lie and φ is not *sd–strategy-proof*. Therefore, for all $\mu \in \varphi(R'_{m_1}, R_{-m_1})$, $\mu(w_1) = m_1$.

Suppose next there is $\mu \in \varphi(R'_{m_1}, R_{-m_1})$ such that $\mu(m_3) \neq w_3$ (i.e., m_3 and w_3 are both single). If m_3 reports R''_{m_3}, as φ is *sd–strategy-proof*, there is $\mu \in \varphi(R'_{m_1}, R''_{m_3}, R_{-\{m_1,m_3\}})$ such that $\mu(m_3) \neq w_3$. If w_3 reports R''_{w_3}, as φ *respects mutual best*, for each $\mu \in \varphi(R'_{m_1}, R''_{m_3}, R''_{w_3}, R_{-\{m_1,m_3,w_3\}})$, $\mu(w_3) = m_3$. Thus, w_3 is better off by telling a lie and φ is not *sd–strategy-proof*. Therefore, for all $\mu \in \varphi(R'_{m_1}, R_{-m_1})$, $\mu(w_3) = m_3$.

We have now pinned down $\varphi(R'_{m_1}, R_{-m_1}) = [\mu_1]$. Analogously, $\varphi(R'_{w_1}, R_{-w_1}) = [\mu_2]$. If $\varphi(R) \neq [\mu_1]$, then m_1 manipulates φ at R by reporting R'_{m_1}. If $\varphi(R) \neq [\mu_2]$, then w_1 manipulates φ at R by reporting R'_{w_1}. Therefore, φ cannot be *sd–strategy-proof*. ∘

We summarize these findings in the following impossibility theorem.

Theorem 1. If a rule is *stable at all times*, then it is *manipulable to the top*. An *sd–strategy-proof* rule does not *respect mutual best*.

4.2. The Compromises and Rewards rule

In this subsection, we design a *weakly group sd–strategy-proof* rule that selects stable sequences. Example 5 explains how the algorithm which defines the rule is designed.

Example 5: Introducing Algorithm 1. For concreteness, consider the two-sided problem with agents $M = \{m_1, m_2, m_3, m_4\}$ and $W = \{w_1, w_2, w_3, w_4\}$ with preferences in Table 4.

R_{m_1}	R_{m_2}	R_{m_3}	R_{m_4}	R_{w_1}	R_{w_2}	R_{w_3}	R_{w_4}
w_1	w_2	w_1	w_4	m_2	m_1	m_3	m_3
		w_3					$\emptyset$

Table 4: Preferences for Example 5. Whenever only partial preferences are provided, missing agents can be ranked arbitrarily below the provided ones.

We will construct a sequence $[\mu_1, \mu_2]$. As a first step, we create a directed graph; see the leftmost graph in Figure 2. In it, each agent is represented by a vertex and has exactly one outgoing edge, namely to the agent's most preferred agent. Here, as w_1 is m_1's most preferred agent, there is an edge (m_1, w_1).

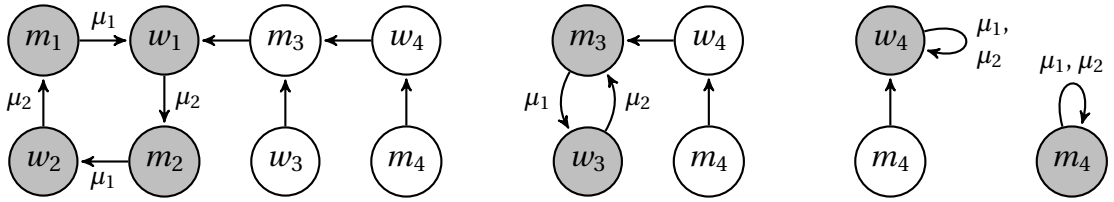

Figure 2: The graphs referred to in Example 5.

The graph must contain either a loop or a cycle as each agent has an outgoing edge and there is a finite set of agents. For the two-sided problem, there can be no cycles of odd length as each cycle must alternate between a member of M and a member of W. Here, the unique cycle $C = (m_1, w_1, m_2, w_2)$ is marked in gray. At μ_1, we match the first agent of the cycle to the one she is pointing to (the second); the third agent to the fourth; and so on. At μ_2, we match the second agent with the third; the fourth with the fifth; ...; and the mth with the first. In other words, agents in C are set to alternate between their neighbours as indicated by the matchings along the edges in Figure 2. The agents in C are then removed and a new graph is created for $N \setminus C$. It contains the cycle $C' = (m_3, w_3)$. We match m_3 and w_3 at μ_1 and at μ_2.[34] The agents in C' are then removed and a new graph is created

[34] Formally, m_3 alternates between his neighbours in the cycle, which "both" are w_3.

for m_4 and w_4. In it, there is a loop (w_4). Here, w_4 is set to be single at μ_1 and at μ_2. Finally, m_4 is the only remaining agent and is also set to be single all the time. Therefore, $\mu_1 = \{(m_1, w_1), (m_2, w_2), (m_3, w_3), (m_4), (w_4)\}$ and $\mu_2 = \{(m_1, w_2), (m_2, w_1), (m_3, w_3), (m_4), (w_4)\}$, and $CR(R) = [\mu_1, \mu_2]$. ∘

Compromises and Rewards. For each $R \in \mathscr{R}^n$, $CR(R)$ is the sequence $[\mu_1, \mu_2]$ obtained when applying Algorithm 1 to R.

Algorithm 1: *Compromises and Rewards*

1. Define the vertex set $V := N$
2. As long as V is not empty,
3. Define the edge set $E := \{(i, j) \in V \times V : \forall k \in V,\ j\ R_i\ k\}$
4. Find a cycle C in the directed graph (V, E), label it $C = (c_1, c_2, \ldots, c_m)$
5a. If $m = 1$, set $\mu_1(c_1) = \mu_2(c_1) = c_1$
5b. If $m = 2$, set $\mu_1(c_1) = \mu_2(c_1) = c_2$
5c. If $m > 2$, set
$\mu_1(c_1) = c_2,\ \mu_1(c_3) = c_4,\ \ldots,\ \mu_1(c_{m-1}) = c_m$
$\mu_2(c_2) = c_3,\ \mu_2(c_4) = c_5,\ \ldots,\ \mu_2(c_m) = c_1$
6. Set $V := V \setminus C$ and repeat Step 2

We are now ready to state the main result on two-sided problems. The proof can be found in the Appendix.

Theorem 2. For two-sided pairing problems $R \in \mathscr{R}^n$, the *Compromises and Rewards* rule is *sd-stable, weakly group sd–strategy-proof*, *side-neutral*, and *anonymous*.

An auxilliary result is Proposition 2 (see Appendix); its main implication is that Algorithm 1 is "path-independent". To be more precise, the sequence obtained is invariant to the order in which the cycles are processed. This is something of a relief: we pointed out that an issue with the *DA* mechanism is that the potentially arbitrary choice of proposer has severe welfare implications. Here, a similarly arbitrary choice *does not* have any implications for sequence. Moreover, this allows us to impose any heuristic we wish for the order in which cycles are chosen. For instance, for computational reasons it might be desirable to prioritize longer cycles over shorter ones.

The *CR* rule is not *individually rational at all times*. Hence, occasionally there is an agent i who is matched to an agent j such that $i\ P_i\ j$. In some settings, this may not be feasible. Modifying Algorithm 1 by requiring that agent i points to her most preferred available agent *among those who prefer i to being single* is not a good way of dealing with this issue. It is easy to construct examples in which one can manipulate the rule by declaring some acceptable agent unacceptable. Assuming $j\ R_j\ i$ whenever $i\ R_i\ j$ only worsens the issue as the preference domain then no longer is "rectangular" (i's domain of preference should be independent of j's reported preference). Strategy-proofness is then not well-defined. Suppose instead we interpret i and j being impossible to match as something predetermined outside the model. Then

agents are only able to misreport their preference regarding acceptable agents. This requires potentially different preference domains $\mathscr{R}_i$ for each $i \in N$ such that $i\ R_i\ j$ implies $i\ R_i'\ j$ for all $R_i' \in \mathscr{R}_i$ and $j\ R_j'\ i$ for all $R_j' \in \mathscr{R}_j$. Our positive findings, Theorems 2, 4, and 5, hold also when we modify the model and consider only problems in $(\mathscr{R}_i)_{i\in N} \subset \mathscr{R}^n$ restricted like this.

5. A short look at many-to-one matching

The *many-to-one matching problem* is one where agents on one side can match to many agents on the other. Formally, a problem is described by (F, W, R, q), where $q = (q_f)_{f\in F}$ is a vector of capacities. For concreteness, think of $f \in F$ as a firm that hires $q_f \in \mathbb{N}$ workers among those in W. We can embed the problem in the one-to-one setting as follows. Each $f \in F$ with capacity q_f is represented by the agents $f^1, f^2, \ldots, f^{q_f}$, each of them having the same preference over W as f has. That is, $R_{f^1} = R_{f^2} = \cdots = R_{f^{q_f}} = R_f$. Moreover, $w \in W$ prefers $f^k \in F$ to $g^m \in F$ whenever $f\ P_w\ g$ or, in case $f = g$, when $k < m$. It is then straightforward to apply Algorithm 1 to this problem. The *CR* rule is still *sd-stable*; the proof is essentially the same as in Theorem 2. However, as a firm can be part of multiple cycles, it may be able to manipulate the rule.

Example 6: Many-to-one manipulation of *CR*. Consider the many-to-one matching problem with agents $F = \{f_1, f_2, f_3\}$ and $W = \{w_1, w_2, \ldots, w_5\}$, quotas $q_{f_1} = 2$, $q_{f_2} = q_{f_3} = 1$, and preferences in Table 5.

$\tilde{R}_{f_1}$	R_{f_1}	R_{f_2}	R_{f_3}	R_{w_1}	R_{w_2}	R_{w_3}	R_{w_4}	R_{w_5}
w_1	w_2	w_1	w_4	f_1	f_2	f_3	f_1	f_1
w_2	w_1	w_5					f_3	
	w_3							
	w_5							
	w_4							

Table 5: Preferences for Example 6. The preference $\tilde{R}_{f_1}$ is reported by f_1 to manipulate the *CR* rule.

The *CR* rule selects $[\mu_1, \mu_2]$, where

$$\mu_1 = \{(f_1^1, w_1), (f_1^2, w_4), (f_2, w_2), (f_3, w_3), (w_5)\}$$
$$\mu_2 = \{(f_1^1, w_2), (f_1^2, w_3), (f_2, w_1), (f_3, w_4), (w_5)\}.$$

Consider the joint manipulation by f_1^1 and f_1^2 where they rank w_1 over w_2 at the top of their preference lists; that is, they report $\tilde{R}_{f_1}$. The resulting sequence is $[\mu_1', \mu_2']$.

$$\mu_1' = \{(f_1^1, w_1), (f_1^2, w_5), (f_2, w_2), (f_3, w_4), (w_3)\}$$
$$\mu_2' = \{(f_1^1, w_1), (f_1^2, w_2), (f_2, w_5), (f_3, w_4), (w_3)\}$$

This is an improvement for f_1^2, though not for f_1^1. That is to be expected as *CR* is *weakly group sd–strategy-proof*. However, it is an improvement for f_1 *as a whole*. When reporting its

true preference, f_1 matches to $\{w_1, w_4\}$ and $\{w_2, w_3\}$. When misreporting its preference, f_1 matches to $\{w_1, w_5\}$ and $\{w_1, w_2\}$. ∘

6. Extending to the general pairing problem

For general pairing problems, there is a set of **agents** N. In contrast to two-sided problems, the agents are not divided into two groups. A (general) **matching** is a mapping $\mu\colon N \to N$ such that $\mu(i) = j \Leftrightarrow \mu(j) = i$ for all $\{i, j\} \subseteq N$. The **set of general matchings** is $\mathscr{M}^*$. All concepts defined regarding matchings in Section 2 can be redefined with respect to $\mathscr{M}^*$ rather than $\mathscr{M}$. Previously, a preference was restricted in the sense that M-agents preferred W-agents to other M-agents. We remove this restriction and denote the **set of general preferences** $\mathscr{R}^*$. A **profile** is $R \in (\mathscr{R}^*)^n$. The **set of sequences of general matchings** is $\mathscr{S}^*$. We extend all concepts previously defined by substituting $\mathscr{S}^*$ for $\mathscr{S}$. This is straightforward as no concepts besides matchings, preferences, and the two permutation axioms refer to the sets M and W. *Side-neutrality* is not applicable for the general problem. A rule is **anonymous** (in welfare terms) if it is invariant to any permutation $\pi \in \Pi$:

$$\forall R \in (\mathscr{R}^*)^n,\ \forall i \in N,\ \forall \pi \in \Pi,\ \varphi(\pi \circ R)\ I_i^{\mathrm{sd}}\ \pi \circ \varphi(R).$$

In contrast to two-sided problems, general problems need not have stable matchings. The problem examined in Example 7 has no stable matching, but does have stable sequences.

Example 7: A general problem without a stable matching. Consider the general problem with agents $N = \{1, 2, 3\}$ with preferences in Table 6.

R_1	R_2	R_3
2	3	1
3	1	2
∅	∅	∅

Table 6: Preferences for Example 7.

For each $\mu \in \mathscr{M}^*$, at least one agent i is single. Then i and her second most preferred agent block μ. Hence, there is no stable matching. In contrast, $\Sigma = [\mu_1, \mu_2, \mu_3]$ is a stable sequence, where $\mu_1 = \big\{(1), (2,3)\big\}$, $\mu_2 = \big\{(1,3), (2)\big\}$, and $\mu_3 = \big\{(1,2), (3)\big\}$.

Note that Σ is "minimal" in the sense that if a matching is removed, then the sequence is no longer stable. For instance, $\{1, 3\}$ block $[\mu_1, \mu_2]$. Intuitively, agent 1 can "guarrantee" himself agent 3 as 3 is always willing to block through $[\mu_2]$. At μ_1, agent 1 therefore makes a compromise by being single. However, agent 1 is not rewarded for this at $[\mu_1, \mu_2]$, in the sense that 1 does not get to match with 2. Therefore, 1 and 3 block $[\mu_1, \mu_2]$. In contrast, Σ is a sequence that balances the "compromises" and "rewards" for all agents.[35] ∘

[35] An immediate consequence is the following. For general pairing problems, if a rule φ is *sd–3-stable*, then $\#\varphi(R) > 2$ for some $R \in (\mathscr{R}^*)^n$. This is different from what we found in Section 4, where we presented two *sd-stable* rules for two-sided problems that never selected sequences of more than two matchings.

6.1. The (Extended) All-Proposing Deferred Acceptance rule

As the agents no longer are divided into two groups, it is not clear how to use the *DA* mechanism. We therefore propose a modified mechanism to extend *RDAAP*. The rule is defined on a restricted domain. The problem in Example 7 is for instance not covered. In contrast, all two-sided problems are. Formally, we define its domain $\mathscr{E}$ as follows. For $R \in (\mathscr{R}^*)^n$, $R \in \mathscr{E}$ if and only if, when applying Algorithm 2 to R, every cycle at Step 4 is of even length.

All-Proposing *DA*. For each $R \in \mathscr{E}$, $APDA(R)$ is the sequence $[\mu_1, \mu_2]$ obtained when applying Algorithm 2 to R.

Algorithm 2: *All-Proposing Deferred Acceptance*

---

1. Create a graph with vertex set $V := N$. An edge (i, j) in the edge set E is associated with the proposal from agent i to j. Initially, each agent proposes (adds an edge in E) to her most preferred agent.

2. Each agent i rejects all proposals but the one made by her most preferred agent j among those proposing (pointing) to her. If $i\ P_i\ j$, then j's proposal to i is also rejected. If a proposal is rejected, its associated edge is removed from E. If no proposals are rejected, continue to Step 3. Otherwise, each rejected agent proposes (add an edge) to her most preferred agent that has not yet rejected her, and Step 2 is repeated.

3. For each cycle C in the directed graph (V, E), labeled $C = (c_1, c_2, \ldots, c_m)$,

4a. If $m = 1$, set $\mu_1(c_1) = \mu_2(c_1) = c_1$

4b. If $m = 2$, set $\mu_1(c_1) = \mu_2(c_1) = c_2$

4c. If $m > 2$, set

$$\mu_1(c_1) = c_2,\ \mu_1(c_3) = c_4,\ \ldots,\ \mu_1(c_{m-1}) = c_m$$
$$\mu_2(c_2) = c_3,\ \mu_2(c_4) = c_5,\ \ldots,\ \mu_2(c_m) = c_1$$

For the rule to be well-defined, each agent has to be part of exactly one cycle at Step 3. If an agent is part of two or more cycles, then some agent must have multiple outgoing edges. But that cannot happen, as the agent only makes a new proposals when her previous one is rejected. If an agent is not part of any cycles, then some agent must have multiple incoming edges. But that cannot happen, as the agent always rejects all but at most one proposal. Hence, the rule is well-defined.

Remark 1: Relation to stable matchings and partitions. The domain $\mathscr{E}$ is logically independent of the set of problems that have stable matchings. The following six-agent examples also show that the graph in Algorithm 2 is fundamentally different from those that arise from Tan's (1991) *stable partitions* (see Appendix C).

(i) Preferences in Table 7*(i)*. The matching $\mu = \{(1,2), (3,4), (5,6)\}$ is stable. The problem is not in $\mathscr{E}$ but in $\hat{\mathscr{E}}$.

(ii) Preferences in Table 7*(ii)*. There is no stable matching. The problem is in $\mathscr{E}$. The *APDA* rule selects $[\{(1,2), (3,4), (5,6)\}, \{(1,6), (2,3), (4,5)\}]$.

The *APDA* rule is *sd-stable*, *anonymous*, and *individually rational at all times* for problems in $\mathscr{E} \subset (\mathscr{R}^*)^n$ (proof is given for a more general result in Theorem 4).[36] We can easily extend

[36] The domain is not rectangular. *Strategy-proofness* is therefore not well-defined.

R_1	R_2	R_3	R_4	R_5	R_6
2	3	1	5	6	4
	1	4	3		5

Table 7*(i)*: Preferences for Remark 1 *(i)*.

R_1	R_2	R_3	R_4	R_5	R_6
2	3	4	5	6	1
3	1	1	6	4	4
6		2	3		5

Table 7*(ii)*: Preferences for Remark 1 *(ii)*.

APDA to a larger domain of problems. Define $\hat{\mathscr{E}}$ by including all problems that have stable matchings, so $\hat{\mathscr{E}} = \mathscr{E} \cup \{R \in (\mathscr{R}^*)^n : \mathscr{C}(R) \neq \emptyset\}$. Let $C(R)$ denote the sequence that attributes equal weight to each stable matching, that is, $C(R) = [\mu : \mu \in \mathscr{C}(R)]$.

Extended *APDA*. For each $R \in \hat{\mathscr{E}}$,

$$EAPDA(R) = \begin{cases} APDA(R) & \text{if } R \in \mathscr{E}, \\ C(R) & \text{otherwise.} \end{cases}$$

This extension is motivated by the following result. It extends Manjunath's (2014) Proposition 3 from two-sided to general pairing problems (using a similar proof technique).

Theorem 3. For general pairing problems $R \in (\mathscr{R}^*)^n$, a sequence of stable matchings is stable.

Theorem 4. For general pairing problems $R \in \hat{\mathscr{E}}$, the *Extended All-Proposing Deferred Acceptance* rule is *sd-stable, anonymous,* and *individually rational at all times.*

6.2. The Generalized Compromises and Rewards rule

Next we present a rule defined on the full preference domain $(\mathscr{R}^*)^n$. Algorithm 3, that defines the rule, coincides with Algorithm 1 in special cases, for instance for two-sided problems. For general pairing problems, it extends Algorithm 1 by being able to handle a certain type of odd cycles.

Example 8: Introducing Algorithm 3. To illustrate the algorithm, consider the general problem with agents $N = \{1, 2, \ldots, 6\}$ with preferences in Table 8. The problem has no stable matching and is not in the domain of *APDA*.

R_1	R_2	R_3	R_4	R_5	R_6
2	3	1	5	2	1
3	1	2	1	1	3
4	5	4	6	6	4
6					

Table 8: Preferences for Example 8.

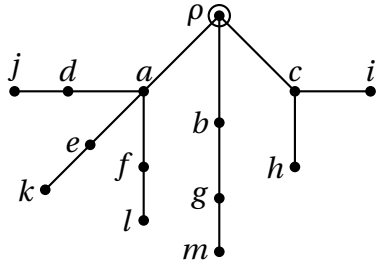

Figure 3: The tree T referred to in Example 8.

Node	Pairs	Weight	Node	Pairs	Weight	Branch β	Matching μ_β	Weight ω_β
ρ	$\emptyset$	1	g	$\{(2,5)\}$	1	$1:\rho, a, d, j$	$\{(1,6),(2,3),(4,5)\}$	1/9
a	$\{(2,3)\}$	1/3	h	$\{(3,6),(4,5)\}$	1/2	$2:\rho, a, e, k$	$\{(1,5),(2,3),(4,6)\}$	1/9
b	$\{(1,3)\}$	1/3	i	$\{(3,4),(5,6)\}$	1/2	$3:\rho, a, f, l$	$\{(1,4),(2,3),(5,6)\}$	1/9
c	$\{(1,2)\}$	1/3	j	$\{(1,6)\}$	1	$4:\rho, b, g, m$	$\{(1,3),(2,5),(4,6)\}$	1/3
d	$\{(4,5)\}$	1/3	k	$\{(4,6)\}$	1	$5:\rho, c, h$	$\{(1,2),(3,6),(4,5)\}$	1/6
e	$\{(1,5)\}$	1/3	l	$\{(5,6)\}$	1	$6:\rho, c, i$	$\{(1,2),(3,4),(5,6)\}$	1/6
f	$\{(1,4)\}$	1/3	m	$\{(4,6)\}$	1			

Table 10: On the left are the nodes of the tree T referred to in Example 8 (Figure 3) with their associated pairs and weights; on the right are the branches of T with their associated matchings and weights.

To keep track of who's matched with whom, we use a tree T that initially only contains its root ρ (Figure 3). To each node of T, we associate a list of pairs and a weight (for the final full list, see Table 10). An agent is *available at node* v if she is not paired along the path from ρ to v. Collect the agents available at v in V_v; note that $V_\rho = N$. If $V_v \neq \emptyset$, we find children of v as follows. Create a directed graph in which each agent in V_v points to her most preferred agent in V_v. In the graph, select a cycle C. Here, at ρ, the graph is the leftmost of Figure 4. It has a unique cycle $C = (1, 2, 3)$.

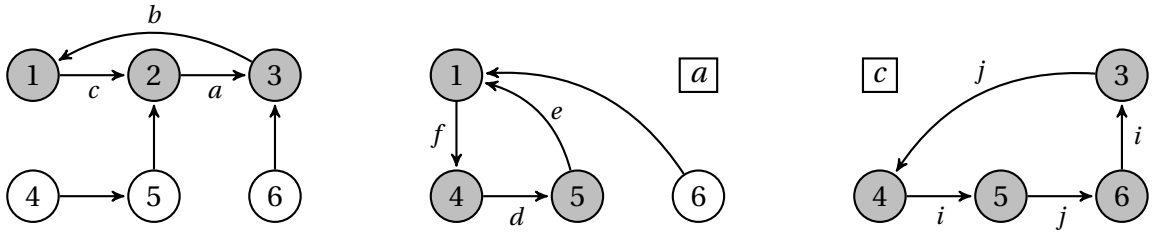

Figure 4: The three graphs referred to in Example 8.

In general, if we select an odd cycle of length m, we add m children of v. At the kth child, the kth agent of C is *not* matched to someone else in C. Here, we add nodes a (where agents 2 and 3 are matched), b (1 and 3), and c (1 and 2). A node's weight is one over its number of siblings; a, b, and c have weight 1/3. At a, the graph is at the middle of Figure 4 and has an odd cycle $(1, 4, 5)$. We pair agents 4 and 5 at node d, 1 and 5 at e, and 1 and 4 at f. Each of these nodes has weight 1/3. At c, the graph is the rightmost of Figure 4 and has an even cycle $(3, 4, 5, 6)$. When we select an even cycle of at least four agents, we add two children. We pair agents 3 and 6, 4 and 5 at node i; agents 3 and 4, 5 and 6 at node j. Each of these nodes has weight 1/2. When we select a cycle of at most two agents, we add one child.

A *leaf* λ of T is a node at which all agents are matched, $V_\lambda = \emptyset$. A *branch* is a collection of connected nodes, starting at the root and ending at a leaf without repeating any nodes. To each branch β, we associate a matching μ_β and a weight ω_β. For instance, branch 1 through a and d to j corresponds to the matching $\mu_1 = \{(1,6),(2,3),(4,5)\}$. The weight is the product of the nodes' weights. For instance, branch 5 through c to h has weight $1 \cdot 1/3 \cdot 1/2 = 1/6$. Finally, find the smallest integer $x \geq 0$ such that, for each branch, the product of x and the branch's

weight is integer. Here, $x = 18$. For each branch β, μ_β is included $\omega_\beta \cdot x$ times in the sequence. For this problem R, we have

$$GCR(R) = [\,\mu_1\,\mu_1,\,\mu_2,\,\mu_2,\,\mu_3,\,\mu_3,\,\mu_4,\,\mu_4,\,\mu_4,\,\mu_4,\,\mu_4,\,\mu_4,\,\mu_5,\,\mu_5,\,\mu_5,\,\mu_6,\,\mu_6\,] \qquad \circ$$

Generalized Compromised and Rewards. For each $R \in (\mathscr{R}^*)^n$, $GCR(R)$ is the sequence obtained when applying Algorithm 3 to R.

Algorithm 3: *Generalized Compromises and Rewards*

1. Initialize a tree T rooted at ρ
2. Use Function 1 to add children of ρ, children of the children of ρ, ...
3. Denote the branches of T by B. For each $\beta \in B$, union the pairs along β to get the graph of a matching μ_β. Its weight ω_β is the product of its nodes' weights.
4. Define $x > 0$ as the smallest integer such that, for each $\beta \in B$, $\omega_\beta \cdot x$ is integer
5. Define $GCR(R)$ as the sequence that, for each $\beta \in B$, contains μ_β exactly $\omega_\beta \cdot x$ times and no other matching

Function 1: Add children of node v

1. Define the vertex set V_v as all agents *not* paired along the path from ρ to v and define the edge set $E_v := \{(i,j) \in V_v \times V_v : \forall\, k \in V_v,\ j\ R_i\ k\}$
2. Find a cycle C in the directed graph (V,E), label it $C = (c_1, c_2, \ldots, c_m)$
3a. If $m = 1$, add one child of v with weight 1. Associate with it (c_1).
3b. If $m = 2$, add one child of v with weight 1. Associate with it (c_1, c_2).
3c. If $m > 2$ is even, add two children of v with weights $1/2$
Associate with the kth (c_k, c_{k+1}), (c_{k+2}, c_{k+3}), ..., (c_{k-1}, c_k)
3d. If $m > 2$ is odd, add m children of v with weights $1/m$
Associate with the kth (c_{k+1}, c_{k+2}), (c_{k+3}, c_{k+4}), ..., (c_{k-2}, c_{k-1})

The further down the tree we move, the smaller the problem we have to solve. Hence, it is clear that Algorithm 3 terminates in finite time. Because an agent is available at a node if and only if she has not yet been matched, the procedure generates well-defined matchings and the rule selects a well-defined sequence. All statements made in Proposition 2 are valid for Algorithm 3. Hence, the order in which cycles are processed is irrelevant for the sequence ultimately chosen. For computational reasons, it is in favorable to process even cycles before odd ones.

Theorem 5. For general pairing problems $R \in (\mathscr{R}^*)^n$, the *Generalized Compromises and Rewards* rule is *sd–5-stable, weakly sd–strategy-proof,* and *anonymous.*

In addition, one can easily show that *GCR* always selects sequences of Pareto-efficient matchings. However, this does not necessarily imply that the sequence is sd-efficient. See Example 11 in the Appendix, building on a similar finding by Bogomolnaia and Moulin (2001). We do however conjecture that we can strengthen the two first properties as follows.

Conjecture 1. For general pairing problems $R \in (\mathscr{R}^*)^n$, the *Generalized Compromises and Rewards* rule is *sd-stable* and *weakly group sd–strategy-proof*.

7. Discussion

7.1. Probabilistic and fractional matchings

A sequence can be reinterpreted as a lottery over matchings (a *probabilistic* matching). The support of the lottery coincides with the set of matchings included in the sequence. The probability assigned to a matching in the lottery is the frequency in which the matching appears in the sequence. For instance, $\Sigma = [\mu, \mu, \mu']$ corresponds to the lottery that assigns probability $2/3$ to μ and $1/3$ to μ'.

The relation between probabilistic and *fractional* matchings is most apparent when we examine the matrix σ that is induced by the sequence Σ. Formally, σ is a symmetric doubly stochastic matrix. Indeed, this is all that is required for σ to represent a fractional matching. For probabilistic matchings, there is an additional requirement for odd groups of agents. For convenience, define $\mathscr{O} = \{S \subseteq N : \#S \geq 3,\ \#S \text{ is odd}\}$.

Fractional and probabilistic matchings. A *fractional matching* f satisfies, for all $\{i, j\} \subseteq N$, $f_{ij} = f_{ji} \in [0, 1]$, and for each $i \in N$, $\sum_{j \in N} f_{ij} \leq 1$. A *probabilistic matching* σ is a fractional matching such that, for each $S \in \mathscr{O}$, $\sum_{i \in S} \sum_{j \in S \setminus \{i\}} \sigma_{ij} \leq \#S - 1$.

Denote the **set of fractional matchings** by $\mathscr{F}^*$. In special cases, for instance for all two-sided problems, fractional matchings coincide with probabilistic matchings (Birkhoff, 1946; von Neumann, 1953). For more recent work on fractional matchings, see Abeledo and Rothblum (1994), Biró and Fleiner (2010), Chiappori et al. (2014), Manjunath (2014), and Budish et al. (2013). An analysis of probabilistic matchings in the two-sided problem can be found in Manjunath (2013). The following two applications highlight the difference between fractional and probabilistic matchings.

Example 9: Fractional and probabilistic matchings. Suppose agents 1, 2, and 3 are hired to work on a project. For the first part of the project, 1 and 2's skills are needed; for the second, 1 and 3's; for the third, 2 and 3's. The fractional matching f pairs 1 half the time with 2, 1 half the time with 3, and 2 half the time with 3. Note that f cannot be interpreted as a lottery over matchings. During the time that 2 and 3 are matched, we require that 1 is single. However, $f_{23} > 0$ and $f_{11} = 0$.

$$f = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix} \qquad \sigma = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

Next, 1, 2, and 3 are having a break from work and consider playing chess. To find out who will play against whom, they draw straws. If 1 and 2 end up playing (being paired), 3 is single. The expected outcome of this lottery is described by σ. This is the probabilistic matching familiar from Example 7. ∘

It should be clear how to extend preferences over sequences (probabilistic matchings) to preferences over fractional matchings. It is equally clear how we can extend blocking, stability, and strategy-proofness to fractional matchings. The rules that we have defined on restricted domains are interesting to study on the full domain if we allow for fractional matchings. Recall, the *CR* rule is a mapping from $\mathscr{R}^n$ to $\mathscr{S}$. Let us define the *Fractional CR* rule, *FCR* for short, on the full domain $(\mathscr{R}^*)^n$. For some problems, *FCR* selects a fractional matching. Hence, the range has changed from $\mathscr{S}$ to $\mathscr{F}^*$.

Fractional *CR*. For each $R \in (\mathscr{R}^*)^n$, $FCR(R)$ is the fractional matching obtained when applying Algorithm 1 to R.

Likewise, we can change the domain and range of *APDA* from $\mathscr{E} \to \mathscr{S}^*$ to the more general $(\mathscr{R}^*)^n \to \mathscr{F}^*$. This yields the *Fractional APDA* rule. For problems not in $\mathscr{E}$, this rule selects a fractional but not probabilistic matching.

Fractional *APDA*. For each $R \in (\mathscr{R}^*)^n$, $FAPDA(R)$ is the fractional matching obtained when applying Algorithm 2 to R.

Theorem 6. For general pairing problems $R \in (\mathscr{R}^*)^n$, the *Fractional All-Proposing Deferred Acceptance* rule is *sd-stable*. On the same domain, the *Fractional Compromises and Rewards* rule is *sd-stable* and *sd–strategy-proof*.

7.2. Completing the preference: expected utility preferences

In this subsection, we define preferences over sequences that are complete. Importantly, for some of these, our main results still hold.

The **utility function** $u_i \colon N \to \mathbb{R}_+$ **represents** $R_i \in \mathscr{R}^*$ if, for each $\{i, j, k\} \subseteq N$, $j \; R_i \; k \Leftrightarrow u_i(j) \geq u_i(k)$. A **profile of utility functions** is $u \equiv (u_i)_{i \in N}$. For each $R \in (\mathscr{R}^*)^n$, let $\mathscr{U}(R)$ denote the collection of profiles of utility functions such that, for each $i \in N$, u_i represents R_i. At $u \in \mathscr{U}(R)$, the **expected utility** for $i \in N$ of $\Sigma \in \mathscr{S}^*$ is

$$U_i(\Sigma) = \sum_{j \in N} \sigma_{ij} u_i(j).$$

For $R \in (\mathscr{R}^*)^n$ and $u \in \mathscr{U}(R)$, define the binary relation R_i^u on $\mathscr{S}^*$ such that, for each $\{\Sigma, \Psi\} \subseteq \mathscr{S}^*$,

$$\Sigma \; R_i^u \; \Psi \Leftrightarrow U_i(\Sigma) \geq U_i(\Psi).$$

In contrast to R_i^{sd}, the relation R_i^u is complete. Define ***u*-blocking** groups and ***u*-stable** sequences by replacing R_i^{sd} by R_i^u in the definitions of sd-blocking and sd-stability. At $R \in (\mathscr{R}^*)^n$, if $S \subseteq N$ weakly blocks $\Sigma \in \mathscr{S}^*$, then S u-blocks Σ for all $u \in \mathscr{U}(R)$. In addition, if there is $u \in \mathscr{U}(R)$ such that S u-blocks $\Sigma \in \mathscr{S}^*$, then S sd-blocks Σ.

In Example 7, $[\mu_1, \mu_2, \mu_3]$ is u-stable for *some* $u \in \mathscr{U}(R)$. The idea is similar to that contemplated in Example 1. For agents to be willing to compromise, the rewards need to be sufficiently big. If each of them find their most preferred agent much better than their second most preferred, they cannot block the sequence.

Example 10: Continuing Example 7. Sequences only containing μ_1, μ_2, and μ_3 can be represented in the three-dimensional simplex. Figure 5 illustrates $[\mu_1, \mu_2, \mu_3]$ together with the upper and lower contour sets of R_1^{sd} at Σ, $U(R_1^{\text{sd}}, \Sigma)$ and $L(R_1^{\text{sd}}, \Sigma)$, where

$$U(R_1^{\text{sd}}, \Sigma) = \{\Psi \in \mathscr{S} : \Psi \; R_1^{\text{sd}} \; \Sigma\} \text{ and } L(R_1^{\text{sd}}, \Sigma) = \{\Psi \in \mathscr{S} : \Sigma \; R_1^{\text{sd}} \; \Psi\}.$$

The figure also illustrates the *line of indifference* (generally: hyperplanes) that intersect the contour sets for various preferences R_1^u. More specifically, let $u \in \mathscr{U}(R)$ be such that, for each $i \in N$, $u_i(i+1) = 1$, $u_i(i-1) = 1/\alpha \equiv \beta \in (0,1)$, and $u_i(i) = 0$. The higher the value of β, the steeper the line of indifference. Intuitively, what matters most to agent 1 is finding *a* partner. In contrast, a low value indicates that it is crucial that the partner is agent 3.

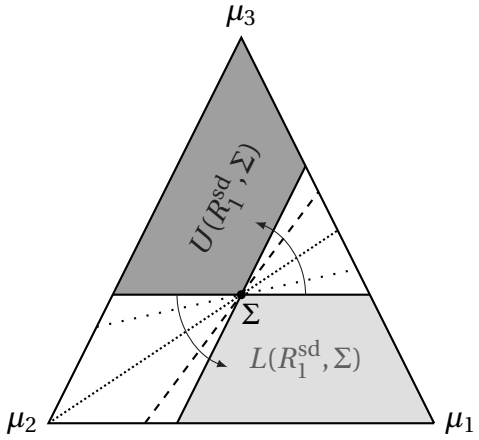

Figure 5: The sequence $\Sigma = [\mu_1, \mu_2, \mu_3]$ is at the center of the simplex. The darker set $U(R_1^{\text{sd}}, \Sigma)$ is the upper contour set of the relation R_1^{sd} at Σ; the lighter is the lower contour set. Sequences in the white areas are not preferred to Σ according to R_1^{sd}, nor is Σ preferred to them. The loosely dotted (densely dotted; dashed) line is the line of indifference for $\beta = 0.2$ ($\beta = 0.5$; $\beta = 0.8$) for R_1^u. The arcs indicate that the higher β, the steeper the line of indifference. The solid lines represent $\beta = 0$ and $\beta = 1$.

So when does there exist a stable sequence? For each $i \in N$, $U_i(\Sigma) = (1+\beta)/3$. Take a generic pair $\{i, i+1\}$. This pair can block to $\Psi = [\mu_{i-1}]$. We have $U_i(\Psi) = \beta$ and $U_{i+1}(\Psi) = 1 > \beta$. The pair therefore blocks if $\beta \geq (1+\beta)/3$, that is, if $\beta \geq 1/2$. Therefore, Σ is stable whenever $\beta < 1/2$. Later, we find that, more importantly, the fraction $u_i(i+1)/u_i(i-1) = 1/\beta = \alpha$ should exceed 2. ∘

In the proofs of our theorems, we do not exploit that the relation R_i^{sd} is incomplete. Neither do we use the discontinuous nature of the relation. Indeed, we can generalize the results by allowing for a larger class of preferences.

***α*-proportional utility functions.** For each $\alpha \geq 1$, each $R \in (\mathscr{R}^*)^n$, and each $u \in \mathscr{U}(R)$,

$$u \in \mathscr{U}_\alpha(R) \Leftrightarrow \forall \{i, j, k\} \subseteq N, \; j \; P_i \; k \Rightarrow u_i(j) > \alpha u_i(k).$$

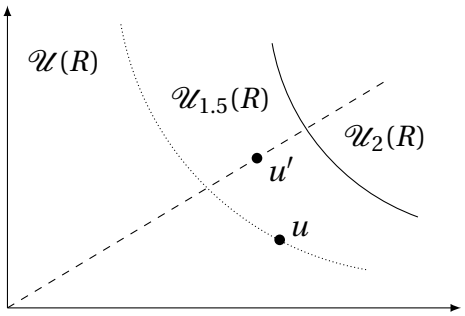

Figure 6: An illustration of some domains of profiles of utility functions for Example 10. The full set is $\mathscr{U}(R)$. The domains are nested: if $\alpha \geq \alpha'$, then $\mathscr{U}_\alpha(R) \subseteq \mathscr{U}_{\alpha'}(R)$. The dashed line contains the profiles of symmetric utility functions examined in Example 10. We argue that the maximal domain with respect to the parameter α for which there always are 2-stable sequences is $\mathscr{U}_2(R)$. This is not to say that this is the maximal domain with respect to set inclusion. For instance, there is a profile $u \in \mathscr{U}_{1.5}(R)$ for which there is a 2-stable sequence. However, for our statement, if we include u, then we have to include $\mathscr{U}_{1.5}(R)$ in its entirety. But then there is $u' \in \mathscr{U}_{1.5}(R)$ for which there is no 2-stable sequence.

Example 10 shows that it is necessary to remove some profiles from $\mathscr{U}(R) \equiv \mathscr{U}_1(R)$ if we wish to find a stable rule – even if it is just *2-stable*. In particular, in terms of the parameter α, there is no domain larger than $\mathscr{U}_2(R)$ for which there always are 2-stable sequences. See Figure 6 for an illustration. As it turns out, this domain is indeed the largest that always have 2-stable sequences. In this way, we can extend our main theorems as follows.

Theorem 1*. For two-sided pairing problems $R \in \mathscr{R}$ with profiles of utility functions $u \in \mathscr{U}_2(R)$, the *Compromises and Rewards* rule is *u-stable* and *weakly group u–strategy-proof*.

Theorem 3*. For general pairing problems $R \in \mathscr{E}^*$ with profiles of utility functions $u \in \mathscr{U}_2(R)$, the *Extended All-Proposing Deferred Acceptance* rule is *u-stable*.

Theorem 4*. For general pairing problems $R \in (\mathscr{R}^*)^n$ with profiles of utility functions $u \in \mathscr{U}_\alpha(R)$ for large enough $\alpha \in \mathbb{R}$, the *Generalized Compromises and Rewards* rule is *u–5-stable* and *u–strategy-proof*.

8. Conclusion

We study two-sided ("marriage") and general pairing ("roommate") problems. We introduce "sequences," lists of matchings that are repeated in order. Stable sequences are natural extensions of stable matchings; case in point, we show that a sequence of stable matchings is stable. In addition, stable sequences can provide solutions to problems for which stable matchings do not exist. In a sense, they allow us to "balance" the interest of the agents at different matchings. In this way, sequences can be superior to matchings in terms of welfare and fairness.

A seminal result due to Roth (1982) is that no *strategy-proof* rule always selects stable matchings. In contrast, we show that there is a *weakly group sd–strategy-proof* rule that selects stable sequences. We call it the *Compromises and Rewards* rule, *CR* for short. We show

that the *CR* rule satisfies two appealing fairness axioms: *anonymity* and *side-neutrality*. For the general problem, the *Generalized CR* rule, *GCR* for short, is *sd–5-stable* (cannot be blocked by groups of five or fewer agents), *weakly sd–strategy-proof*, and *anonymous*. In addition, the *Extended All-Proposing Deferred Acceptance* rule is *sd-stable*, *anonymous*, and *individually rational at all times* on a restricted domain. The domain includes all problems that have stable matchings and some that do not. We provide also two negative findings. Namely, rules that are *stable at all times* are *manipulable to the top*. Moreover, *sd–strategy-proof* rules do not *respect mutual best*.

There are still many open questions that are interesting to study, we list only the ones we feel are most important. Clearly, proving Conjecture 1 is one of them (that is, proving that the *GCR* rule is *sd-stable* and *weakly group sd–strategy-proof*. Another is the question of how to extend the *APDA* rule to be defined on a larger, perhaps even the full, domain. Next is the question of how to generalize and complete the preferences over sequences. In a sense, we already provide one answer to this in Subsection 7.2, though only verbally. Similarly, we describe only a sketch of the non-cooperative foundation for stable sequences; this would be interesting to formalize. Lastly, all of our results are of the style that a particular rule satisfies certain axioms. An interesting question is to detail which axioms our rules are *the only* to satisfy.

References

Abdulkadiroğlu, A., Sönmez, T., 2003. School Choice: A Mechanism Design Approach. American Economic Review 93 (3), 729–747.

Abeledo, H., Rothblum, U. G., 1994. Stable Matchings and Linear Inequalities. Discrete Applied Mathematics 54 (1), 1–27.

Abeledo, H., Rothblum, U. G., 1995. Paths to Marriage Stability. Discrete Applied Mathematics 63 (1), 1–12.

Abraham, D. J., Biró, P., Manlove, D. F., 2006. "Almost Stable" Matchings in the Roommates Problem. Approximation and Online Algorithms: Lecture Notes in Computer Science 3879, 1–14.

Akahoshi, T., 2014. A Necessary and Sufficient Condition for Stable Matching Rules to be Strategy-proof. Social Choice and Welfare 43, 683–702.

Alcalde, J., 1995. Exchange-proofness or Divorce-proofness? Stability in One-sided Matching Markets. Economic Design 1, 275–287.

Alcalde, J., Barberà, S., 1994. Top Dominance and the Possibility of Strategy-proof Stable Solutions to Matching Problems. Economic Theory 4 (3), 417–435.

Alkan, A., Tuncay, A., Apr. 2014. Pairing Games and Markets. Working Papers 2014.48, Fondazione Eni Enrico Mattei.

Andersson, T., Erlanson, A., Gudmundsson, J., Habis, H., Ingebretsen Carlson, J., Kratz, J., 2014a. A Method for Finding the Maximal Set in Excess Demand. Economics Letters 125 (1), 18–20.

Andersson, T., Gudmundsson, J., Talman, D., Yang, Z., 2014b. A Competitive Partnership Formation Process. Games and Economic Behavior 86, 165–177.

Ashlagi, I., Fischer, F., Kash, I. A., Procaccia, A. D., 2013. Mix and Match: A Strategyproof Mechanism for Multi-hospital Kidney Exchange. Games and Economic Behavior, in press.

Birkhoff, G., 1946. Three Observations on Linear Algebra. Revi. Univ. Nac. Tucuman, ser A (5), 147–151.

Biró, P., Fleiner, T., 2010. Fractional Solutions for NTU-games. In Proceedings of COMSOC 2010: 3rd International Workshop on Computational Social Choice, 283–294.

Blumgart, H. L., 1964. Caring for the Patient. New England Journal of Medicine (270), 449–456.

Bogomolnaia, A., Moulin, H., 2001. A New Solution to the Random Assignment Problem. Journal of Economic Theory 100 (2), 295–328.

Bosslet, G. T., Torke, A. M., Hickman, S. E., Terry, C. L., Helft, P. R., 2011. The Patient-Doctor Relationship and Online Social Networks: Results of a National Survey. Journal of General Internal Medicine 26 (10), 1168–1174.

Braun, S., Dwenger, N., Kübler, D., 2010. Telling the Truth May Not Pay Off: An Empirical Study of Centralized University Admissions in Germany. The B.E. Journal of Economic Analysis & Policy 10 (1).

Budish, E., Che, Y.-K., Kojima, F., Milgrom, P., 2013. Designing Random Allocation Mechanisms: Theory and Applications. American Economic Review 103, 585–623.

Chiappori, P.-A., Galichon, A., Salanié, B., 2014. The Roommate Problem - Is More Stable Than You Think. CESifo Working Paper Series 4676, CESifo Group Munich.

Diamantoudi, E., Miyagawa, E., Xue, L., 2004. Random Paths to Stability in the Roommate Problem. Games and Economic Behavior 48, 18–28.

Gale, D., Shapley, L., 1962. College Admissions and the Stability of Marriage. American Mathematical Monthly 69, 9–15.

Griffiths, T., May 2002. The Vigilant Lifeguard. Aquatics International.

Gudmundsson, J., May 2013. Cycles and Third-Party Payments in the Partnership Formation Problem. Working Papers 2013:16, Lund University, Department of Economics.

Gudmundsson, J., 2014. When do Stable Roommate Matchings Exist? A Review. Review of Economic Design 18 (2), 151–161.

Gudmundsson, J., Habis, H., Aug. 2013. Assignment Games with Externalities. Working Papers 2013:27, Lund University, Department of Economics.

Inarra, E., Larrea, C., Molis, E., 2008. Random Paths to P-stability in the Roommate Problem. International Journal of Game Theory 36, 461–471.

Inarra, E., Larrea, C., Molis, E., Feb. 2010. The Stability of the Roommate Problem Revisited. CORE Discussion Papers 2010007, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).
Klaus, B., Klijn, F., Walzl, M., 2010. Stochastic Stability for Roommate Markets. Journal of Economic Theory 145, 2218–2240.
Klaus, B., Klijn, F., Walzl, M., 2011. Farsighted Stability for Roommate Markets. Journal of Economic Theory 13, 921–933.
Laffont, J.-J., Tirole, J., 1993. A Theory of Incentives in Procurement and Regulation. MIT Press.
Manjunath, V., 2013. Stability and the Core of Probabilistic Marriage Problems. Working Paper.
Manjunath, V., 2014. Markets for Fractional Partnerships. Working Paper.
Manlove, D. F., 2013. Algorithmics of Matching under preferences. World Scientific.
Mas, A., Moretti, E., 2009. Peers at Work. American Economic Review 99 (1), 112–145.
Niederle, M., Roth, A., 2003. Unraveling Reduces Mobility in a Labor Market: Gastroenterology with and without a Centralized Match. Journal of Political Economy 111 (6), 1342–1352.
Ostrovsky, M., Schwarz, M., 2010. Information Disclosure and Unraveling in Matching Markets. American Economic Journal: Microeconomics 2 (2), 34–63.
Pais, J., Pintér, A., Veszteg, R. F., 2011. College Admission and the Role of Information: An Experimental Study. International Economic Review 52 (3), 713–737.
Pápai, S., 2004. Unique Stability in Simple Coalition Formation Games. Games and Economic Behavior 48, 337–354.
Pathak, P. A., Sönmez, T., 2008. Leveling the Playing Field: Sincere and Strategic Players in the Boston Mechanism. American Economic Review 98 (4), 1636–1652.
Rodríguez-Álvarez, C., 2009. Strategy-proof Coalition Formation. International Journal of Game Theory 38 (3), 431–452.
Roth, A., 1982. The Economics of Matching: Stability and Incentives. Mathematics of Operations Research 7 (4), 617–628.
Roth, A., 2002. The Economist As Engineer: Game They, Experimentation, and Computation As Tools for Design Economics. Econometrica 70 (4), 1341–1378.
Roth, A., Peranson, E., 1999. The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design. American Economic Review 89 (4), 748–780.
Roth, A., Sönmez, T., Ünver, M. U., 2004. Kidney Exchange. Quarterly Journal of Economics 119 (2), 457.
Roth, A., Sotomayor, M., 1990. Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis. Econometric Society Monographs, Cambridge University Press, New York.
Roth, A., Vande Vate, J. H., 1990. Random Paths to Stability in Two-sided Matching. Econometrica 58, 1475–1480.
Sasaki, H., Toda, M., 1996. Two-Sided Matching Problems with Externalities. Journal of Economic Theory 70 (1), 93–108.
Shapiro, C., Stiglitz, J. E., 1984. Equilibrium Unemployment as a Worker Discipline Device. American Economic Review 74 (3), 433–444.
Shapley, L., Scarf, H., 1974. On Cores and Indivisibility. Journal of Mathematical Economics 1, 23–37.
Shapley, L., Shubik, M., 1971. The Assignment Game I: The Core. International Journal of Game Theory 1, 111–130.
Snyder, L., 2012. American College of Physicians Ethics, Professionalism, and Human Rights Committee. American College of Physicians Ethics Manual: sixth edition. Annals of Internal Medicine 156 (1), 73–104.
Sönmez, T., Ünver, M. U., 2010. Matching, Allocation, and Exchange of Discrete Resources.
Talman, D., Yang, Z., 2011. A Model of Partnership Formation. Journal of Mathematical Economics 47, 206–212.
Tan, J. J. M., 1990. A Maximum Stable Matching for the Roommate Problem. BIT 29, 631–640.
Tan, J. J. M., 1991. A Necessary and Sufficient Condition for the Existence of a Complete Stable Matching. Journal of Algorithms 12, 154–178.
The American Red Cross, 2012. Lifeguarding Manual. Krames StayWell Strategic Partnerships Division.
Thomson, W., December 2013. Strategyproof Allocation Rules. mimeo.
Tirole, J., 1986. Hierarchies and Bureaucracies: On the Role of Collusion in Organizations. Journal of Law, Economics, & Organization 2 (2), 181–214.

von Neumann, J., 1953. A certain zero-sum two-person game equivalent to the optimal assignment problem. Vol. 2. Contributions to the theory of games.
von Neumann, J., Morgenstern, O., 1947. Theory of Games and Economic Behavior. Princeton: Princeton University Press, NJ.

Appendix A. Proofs

Appendix A.1. Proof of Theorem 2

Stability: Let $R \in \mathscr{R}^n$ denote a typical problem. To obtain a contradiction, suppose $\Sigma \equiv CR(R)$ is not stable. Assume $S \subseteq N$ block Σ to Ψ, and that no $T \subset S$ blocks Σ. Let C be the first cycle containing an agent from S, say agent i. Label $C = (i, i+1, \ldots, i+m)$. Define M to contain all agents except those that are part of cycles processed prior to C. For all $j \in M$, $(i+1)\ P_i\ j$. Moreover, no agent in $N \setminus M$ is part of S as i is the first to be. Note that if $\#C \leq 2$, then $\sigma_{i,i+1} = 1$ and i is matched the entire time with her most preferred available agent. Agent i is not able to improve upon that. Additionally, $S \setminus \{i, i+1\} \subset S$ can then also block, a contradiction. For i to block, we require therefore $(i+1) \in S$ and $\psi_{i,i+1} \geq \sigma_{i,i+1} = 1/2$. For $i+1$ to block, we require $(i+2) \in S$ and $\psi_{i+1,i+2} \geq \sigma_{i+1,i+2} = 1/2$. As $\psi_{i,i+1} + \psi_{i+1,i+2} \leq 1$, we have $\psi_{i,i+1} = \psi_{i+1,i+2} = 1/2 = \sigma_{i,i+1} = \sigma_{i+1,i+2}$. Repeat the argument for the rest of C. We find that $C \subseteq S$. Moreover, each agent in C is matched under Ψ as she is under Σ. Then no agent in C is better off. Hence, $S \setminus C$ is matched entirely among themselves and includes someone better off under Ψ. Then $S \setminus C \subset S$ can block, a contradiction as no $T \subset S$ was assumed to be able to block.

Group–strategy-proofness: To obtain a contradiction, suppose the *CR* rule is group-manipulable. Assume $S \subseteq N$ can manipulate *CR* at $R \in \mathscr{R}^n$ through $R'_S \in \mathscr{R}^S$, and that no $T \subset S$ can manipulate. Denote $\Sigma \equiv CR(R)$ and $\Psi \equiv CR(R'_S, R_{-S})$. Let C be the first cycle containing an agent from S, say agent i. Label $C = (i, i+1, \ldots, i+m)$. We also use the labelling $C = (i, i-m \equiv i+1, \ldots, i-1 \equiv i+m)$. Define M to contain all agents except those from cycles processed prior to C. For all $j \in M$, $i+1\ P_i\ j$. From Proposition 2, no matter in which order the cycles are processed, the final outcome is the same. In particular, no matter i's reported preference, each cycle processed prior to C is still a cycle. Agent i can therefore not be matched to an agent in $N \setminus M$. Note that if $\#C \leq 2$, then i is matched the entire time with her most preferred available agent. Agent i is not able to improve upon that. Additionally, $S \setminus \{i, i+1\} \subset S$ can then also manipulate, a contradiction. For i to manipulate, we need therefore $i+1 \in S$ and $\psi_{i,i+1} \geq \sigma_{i,i+1} = 1/2$.

Assume next $S \cap C = \{i\}$, that is, that i is the only manipulating agent in C. Note that, no matter i's reported preference, $i-1$ will point to i as long as i is available. Similarly, $i-2$ will point to $i-1$ as long as $i-1$ is available, which is as long as i is. Repeating the argument, $i+1$ points to $i+2$ as long as i is available. Therefore, no matter i's report, the only way for i to be matched to $i+1$ is by pointing to $i+1$. This completes the cycle C and i is not better off. As above, $S \setminus \{i\} \subset S$ then can manipulate, a contradiction. Hence, there are multiple agents in $S \cap C$.

Recall, for i to manipulate, $\psi_{i,i+1} \geq 1/2$. Suppose i is matched to $i+1$ through i pointing to $i+1$. Note that $i+2$ is $i+1$'s most preferred available agent. Given that $\psi_{i,i+1} \geq 1/2$, $i+1$ can do

no better than reporting his preference truthfully (as then $\psi_{i+1,i+2} = 1/2 = \sigma_{i+1,i+2}$). However, now we can repeat the argument for all agents in C. Each of them is matched under Ψ as under Σ. Then $S \setminus C \subset S$ can manipulate, a contradiction. Hence, for i to manipulate, it must be that i is matched to $i+1$ through $i+1$ pointing to i. Then $i+1$ is part of the manipulating group, hence $i+1 \in S$. As $\sigma_{i+1,i+2} = 1/2$, $i+2$ needs to match to $i+1$ the remaining time. Hence, $i+2$ points to $i+1$, and $i+2 \in S$. When we repeat the argument, we find that all agents in C point in the opposite direction. All of them are matched exactly as if they reported their true preference. Again, $S \setminus C \subset S$ can then manipulate, a contradiction.

Side-neutrality and *anonymity* are both immediate. No agent nor any side receives any "special" treatment. □

Appendix A.2. Proof of Theorem 3

We wish to first show that, for all $\{i, j\} \subseteq N$,

$$\sum_{kP_ij} \sigma_{ik} + \sum_{kP_ji} \sigma_{jk} + \sigma_{ij} \geq 1, \tag{1}$$

where σ is the matrix representation of Σ. Suppose, to obtain a contradiction, for some $\{i, j\} \subseteq N$ the inequality is not true. Then,

$$\sum_{kP_ij} \sigma_{ik} + \sum_{kP_ji} \sigma_{jk} + \sigma_{ij} < 1,$$

and therefore there exists some $\mu \in \Sigma$ such that neither $\mu(i)\ P_i\ j$, $\mu(j)\ P_j\ i$, nor $\mu(i) = j$. But then $\{i, j\}$ block μ, a contradiction as μ is stable.

By contradiction, suppose $S \subseteq N$ block Σ to $\Psi \in \mathscr{S}$ with matrix representation ψ. For all $\{i, j\} \subseteq S$, as $\Psi\ R_i^{\text{sd}}\ \Sigma$ and $\Psi\ R_j^{\text{sd}}\ \Sigma$,

$$\sum_{kP_ij} \psi_{ik} \geq \sum_{kP_ij} \sigma_{ik} \text{ and } \sum_{kP_ji} \psi_{ik} + \psi_{ij} \geq \sum_{kP_ji} \sigma_{ik} + \sigma_{ij}.$$

Adding these inequalities and using (1),

$$\sum_{kP_ij} \psi_{ik} + \sum_{kP_ji} \psi_{ik} + \psi_{ij} \geq \sum_{kP_ij} \sigma_{ik} + \sum_{kP_ji} \sigma_{ik} + \sigma_{ij} \geq 1.$$

Let $i \in S$ be such that $\Psi\ P_i^{\text{sd}}\ \Sigma$. Then $\psi_{ij} > \sigma_{ij}$ for some $j \in S$. Therefore,

$$\sum_{kP_ij} \psi_{ik} + \sum_{kP_ji} \psi_{jk} + \psi_{ij} > \sum_{kP_ij} \psi_{ik} + \sum_{kP_ji} \psi_{jk} + \sigma_{ij} \geq \sum_{kP_ij} \sigma_{ik} + \sum_{kP_ji} \sigma_{jk} + \sigma_{ij} \geq 1 \tag{2}$$

In parallel, consider the problem (S, R_S). For each $\mu \in \Psi$, denote the projection of μ onto S by μ_S. To be more precise, for all $i \in S$, $\mu_S(i) = \mu(i)$. As S block Σ to Ψ, $\mu(i) \in S$ for all $i \in S$. Hence μ_S is a well-defined matching in the problem (S, R_S). Let $\Psi_S \equiv [\mu_S : \mu \in \Psi]$ with matrix representation ψ_S. Note that, for all $\{i, j\} \subseteq S$,

$$\sum_{kP_ij} \psi_{S_{ij}} = \sum_{kP_ij} \psi_{ij} \text{ and } \psi_{S_{ij}} = \psi_{ij}.$$

As $\psi_{ij} > \sigma_{ij} \geq 0$, $\mu(i) = j$ for some $\mu \in \Psi$, and hence $\mu_S(i) = j$ for some $\mu_S \in \Psi_S$. By Lemma 3 applied to (S, R_S) and Ψ_S,

$$\sum_{kP_ij} \psi_{S_{ik}} + \sum_{kP_ji} \psi_{S_{jk}} + \psi_{S_{ij}} = 1,$$

where the left hand side equals

$$\sum_{kP_ij} \psi_{ik} + \sum_{kP_ji} \psi_{jk} + \psi_{ij} = 1,$$

This contradicts (2). Hence, S cannot block Σ, and Σ is therefore stable. □

Appendix A.3. Proof of Theorem 4

Stability: To obtain a contradiction, suppose $S \subseteq N$ at $R \in \mathscr{E}$ block $\Sigma \equiv \text{APDA}(R)$ through $\Psi \in \mathscr{S}^*$. Assume S is minimal, in the sense that no $T \subset S$ can block Σ. Let i be an arbitrary agent in S. Suppose i is part of the cycle C. Label $C = (i, i+1, \ldots, i+m \equiv i-1)$. Hence, $\sigma_{i,i+1} = \sigma_{i,i-1} = 1/2$. For each agent j preferred by i to both $i+1$ and $i-1$, j rejected i's proposal. Moreover, j did not propose to i. Hence, at Σ, j is matched only to agents preferred to i. Therefore i cannot match to j at Ψ (then $j \in S$, but j would not find Ψ at least as good as Σ).

Case 1: Suppose $i+1 \; P_i \; i-1$. For $\Psi \; R_i^{\text{sd}} \; \Sigma$, we require $\psi_{i,i+1} \geq \sigma_{i,i+1} = 1/2$. Therefore $i+1 \in S$. From Lemma 1(i), $i+2 \; P_{i+1} \; i$. As i is chosen arbitrarily, the argument applies to $i+1$ as well. Hence, $i+1$ cannot match to some agent preferred to $i+2$ at Ψ. For $\Psi \; R_{i+1}^{\text{sd}} \; \Sigma$, we require $\psi_{i,i+1} \leq \sigma_{i,i+1}$. Hence, $\psi_{i,i+1} = 1/2$. Moreover, we require $\psi_{i+1,i+2} \geq \sigma_{i+1,i+2} = 1/2$. Repeat the argument for $i+2$. We find that $i+1$ is matched identically under Ψ as under Σ. Moreover, by repeating the argument for each agent in C, we find that $C \subseteq S$. Importantly, each agent in C is matched in the same way under Ψ as under Σ. Hence, no agent in C is better off. Then $S \setminus C \subset S$ can block Σ. This is a contradiction.

Case 2: Suppose $i-1 \; P_i \; i+1$. Instead we make use of part (ii) of Lemma 1. Otherwise, the proof is as in Case 1.

Anonymity and *individual rationality at all times* are immediate. □

Appendix A.4. Proof of Theorem 5

Strategy-proofness: The statement is basically proven by repeating the arguments presented in the section on the two-sided problem. The difference is here that a supposedly manipulating agent i can be part of multiple cycles due to the splits of the procedure, rather than just one. However, for each of those cycles, by the same logic as before, i cannot gain by misreporting his preference.

To obtain a contradiction, say i can manipulate. Let $C = (i, i+1, \ldots, i+m \equiv i-1)$ be an arbitrary cycle that includes i. No matter i's reported preference, all cycles C' processed prior to C are cycles. (We can find a "path" of cycles chosen from the start of the algorithm to the point when C is chosen. It can be helpful to have the picture of the timeline in mind here.) Hence, i cannot match with the agents that are unavailable when C is chosen.

i can misreport his preference in such a way that he is taken as part of a different cycle $D = (i, i'+1, \ldots, i'+m' \equiv i'-1)$ such that $i'-1 \; P_i \; i-1$ and $i+1 \; P_i \; i'+1$. However, to compensate

for this loss of time spent with $i+1$, i would have to be taken as part of a different cycle at a different occasion as well. At that occasion, he would similarly swap the agent pointing to him at the expense of time with his, at the time, most preferred agent. This requires yet a different cycle where i changes. However, as there is a finite number of cycles, i will in the end not be able to make up for the loss of time with his, at the time, most preferred agent. Hence, i cannot manipulate.

2-stability: Let $\hat{\Sigma}$ be the sequence constructed. Clearly, no agent i can block $\hat{\Sigma}$ on her own, as if i always is single she cannot strictly improve upon $\hat{\Sigma}$, and otherwise i must at some point be matched to some j P_i i. To obtain a contradiction, suppose $\{i, j\} \subseteq N$ block $\hat{\Sigma}$ to Ψ. It is without loss to assume $\Psi = [\mu]$ such that $\mu(i) = j$. Then, Ψ R_i^{sd} $\hat{\Sigma}$ implies j R_i $\mu(i)$ for each $\mu \in \hat{\Sigma}$. Moreover, say i is chosen as part of a cycle in the algorithm no later than j is. Hence, j is available at the time. If i points to some $k \neq j$, then l P_i j and there exists $\mu \in \hat{\Sigma}$ such that $\mu(i) = k$ P_i j, a contradiction. Hence, i points to j. Then, i and j are chosen in the same cycle. Applying the same argument to j, if j were to point to someone else than i, we obtain a contradiction. Hence, the cycle chosen contains only i and j. But then the same argument applies to the next time, if any, i is chosen as part of a cycle. Hence, i and j must always be matched, and hence neither i nor j prefers Ψ to $\hat{\Sigma}$. This is a contradiction.

5-stability: By Lemma 2, it is immediate that groups S of size 1, 2, and 3 cannot block. The first cycle needs to contain at least three agents, though there must also be agents in S that are not in the cycle. Suppose $S = \{1, 2, 3, 4\}$ can block, and the "first" cycle is $(1, 2, 3)$. Then 1 still needs to be matched a third with his most preferred agent, 2, when blocking. The same goes for 2 with 3 and 3 with 1. The remaining time 1 needs to be matched with agent 4 when blocking. Hence, 1 cannot be taken in a cycle with agents he prefers to 4. The same goes for agents 2 and 3 in the other splits. Therefore agent 4 cannot be matched with anyone 4 prefers to 1, 2 and 3. But then no one can be strictly better off when blocking. This is a contradiction. The proof that $S = \{1, 2, 3, 4, 5\}$ cannot block is similar, but we need to consider more cases. For now, it is available upon request. □

Appendix B. Additional results

Proposition 1. Let $R \in \mathscr{R}^n$ be a general pairing problem. The sequence $\Sigma \in \mathscr{S}^*$ is 2-stable if and only if there is no $\mu \in \mathscr{M}^*$ and $S \subseteq N$ such that S block Σ through $[\mu]$.

Proof. If S block Σ through $[\mu]$, then there exists $i \in S$ and $j \equiv \mu(i)$ such that $[\mu]$ P_i^{sd} Σ and $[\mu]$ R_j^{sd} Σ. Then $\{i, j\}$ block Σ through $[\mu]$. Hence Σ is not 2-stable.

Assume for each $\mu \in \mathscr{M}^*$, there exists no $S \subseteq N$ that blocks Σ through $[\mu]$. We wish to show that (i) no $\{i\} \subseteq N$ and (ii) no $\{i, j\} \subseteq N$ can block Σ. For case (i), if i blocks through Ψ, then i is single at each matching in Ψ. Let $\mu \in \Psi$ be an arbitrary matching. Then i blocks Σ through $[\mu]$, a contradiction. For case (ii), if $\{i, j\}$ block through Ψ, then either they are either single or matched together at each matching in Ψ. If i P_i j, then i can block on her own, a contradiction. Hence j P_i i. Likewise, we must have i P_j j. But then i and j can block through $[\mu]$ for any $\mu \in \Psi$ such that $\mu(i) = j$. This is a contradiction. □

Proposition 2. Algorithm 1 has the following properties.

(i) At each step, each $i \in N$ is part of at most one cycle.

(ii) If C' is a cycle when cycle C is chosen, C' remains a cycle.

(iii) If D becomes a cycle after C is chosen, and D' becomes a cycle after C' is chosen, then $D \cap D' = \emptyset$.

Proof. (i) If there is $i \in C \cap C'$, then there is $j \in C \cap C'$ such that j is followed by k in C and $k' \neq k$ in C', requiring j to point to both k and k', a contradiction.

(ii) Each $i \in C'$ points to $(i+1) \in C'$ such that $(i+1) \notin C$ as $C \cap C' = \emptyset$ by part (i). Then i points to $i+1$ after agents in C are removed as well.

(iii) First removing C leaves C' by part (ii) and D by assumption. Then removing C' leaves D by part (ii). Switching the order of C and C' leaves D', but the remaining agents are the same, hence D and D' are cycle when $C \cup C'$ are removed. By part (i), $D \cap D' = \emptyset$. □

Lemma 1. Let $C = (1, 2, \ldots, m)$ be a cycle encountered in Algorithm 2. (i) If $2\ P_1\ m$, then $k+1\ P_k\ k-1 \pmod m$ for all $k = 1, 2, \ldots, m$. (ii) If $m\ P_1\ 2$, then $k-1\ P_k\ k+1 \pmod m$ for all $k = 1, 2, \ldots, m$.

Proof. (i) Assume $2\ P_1\ m$. To obtain a contradiction, suppose $1\ P_2\ 3$. Prior to proposing to 3, 2's proposal to 1 was rejected. But then 1 should also have rejected m's proposal. This is a contradiction. Hence, $3\ P_2\ 1$. To complete the proof, apply the same argument to agents $3, 4, \ldots, m$.

(ii) Assume $m\ P_1\ 2$. To obtain a contradiction, suppose $1\ P_m\ m-1$. Prior to proposing to 2, 1's proposal to m was rejected. But then m should also have rejected $m-1$'s proposal. This is a contradiction. Hence, $m-1\ P_m\ 1$. To complete the proof, apply the same argument to agents $m-1, m-2, \ldots, 2$. □

Lemma 2. Suppose $S \subseteq N$ is a minimal group that can block the sequence selected by the General Compromises and Rewards rule. Consider a step of the algorithm where (a) all agents of S are available and (b) the cycle chosen, call it C, includes members of S. Then

- C contains only members of S
- C contains an odd number (≥ 3) of agents
- C does not contain all members of S.

If there has been a split prior to the step, there may be multiple "first" cycles. Then, if C is a cycle as described above for some part of the split and D for another, C and D are agent-disjoint.

Proof. Assume $i \in S$ is a member of C, and i points to j. To obtain a contradiction, suppose $j \notin S$. Then $j\ P_i\ k$ for all $k \in S$. But then i cannot be better off if i has to be matched only within S. This is a contradiction, hence $j \in S$. Now, reapply the argument for j. By the finiteness of N (and hence of S and C), eventually we complete a cycle only within S. That is, C contains only members of S.

To obtain a contradiction, suppose C contains one or two agents. These agents get to match entirely with their most preferred agent of S. They cannot do better when S blocks. Hence, $S \setminus C \subset S$ can block, a contradiction to S being minimal. Suppose instead C is even of length 4 or more. Then each agent in C gets to spend half their time with their most preferred agent of S. When S blocks, each agent in C therefore will be matched in this way. We reach the same contradiction. Therefore, C contains an odd number of agents.

To obtain a contradiction, suppose C contains every member of S. Then each agent in S gets to match $(\#S-1)/2\#S$ with his most preferred agent of S. When blocking, each agent in S needs to be single $1/\#S$ of the time. This cannot be an improvement over the sequence selected by the *GCR* rule. This again is a contradiction, hence C does not contain all members of S.

Finally, suppose C and D share some agent, say i. Then i will point to the same agent in both C and D, say j. This is because j is i's most preferred agent of S. Repeat for j and the rest of the agents of C and D. We reach the conclusion that the cycles coincide if they overlap. Hence, if there are different "first" cycles, then they share no agents. □

Lemma 3. Consider a generalized pairing problem with agents N with preferences $R \in \mathscr{R}^n$. Let $\Sigma \in \mathscr{S}^*$ be such that, for all $\{i, j\} \subseteq N$,

$$\sum_{kP_ij} \sigma_{ik} + \sum_{kP_ji} \sigma_{jk} + \sigma_{ij} \geq 1.$$

Then, for all $\{i, j\} \subseteq N$ such that $\mu(i) = j$ for some $\mu \in \Sigma$,

$$\sum_{kP_ij} \sigma_{ik} + \sum_{kP_ji} \sigma_{jk} + \sigma_{ij} = 1.$$

Proof. The result can be deduced from Theorem 4.5 in Abeledo and Rothblum (1994). □

Proposition 3. The *CR* rule is not *sd–strategy-proof*.

Proof. Consider the two-sided problem with agents $N = \{m_1, m_2, m_3\}$ and $W = \{w_1, w_2, w_3\}$ with preferences in Table B.11.

R'_{m_1}	R_{m_1}	R_{m_2}	R_{m_3}	R_{w_1}	R_{w_2}	R_{w_3}
w_3	w_1	w_2	w_3	m_2	m_1	m_1
	w_3					
	w_2					

Table B.11: Preferences for the example in the proof of Proposition 3.

In the sequence $CR(R)$, m_1 matches half the time with w_1 and half the time with w_2. In the sequence $CR(R'_1, R_{-1})$, m_1 always matches with w_3. Telling the truth therefore is not *better* than telling a lie (though neither is telling a lie better than telling the truth). □

R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
5	5	6	6	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
6	6	5	5				
7	7	8	8				
8	8	7	7				

Table B.12: Preferences for Example 11.

Example 11: A sequence of Pareto-efficient matchings need not be sd-efficient. Here, we show that, if each matching in $\Sigma \in \mathscr{S}^*$ is Pareto-efficient, Σ may not be sd-efficient. The result follows from modifying an example in Bogomolnaia and Moulin (2001). The agents are $N = \{1, 2, \ldots, 8\}$ with preferences in Table B.12. We create a sequence Σ of $4! = 24$ Pareto-efficient matchings, where each matching is associated to an ordering of $\{1, 2, 3, 4\}$. At each matching, align agents 1 through 4 according to the ordering, and let them choose partners sequentially. In the array below, row r refers to agent r; column c refers to agent $4 + c$:

$$\sigma = \begin{pmatrix} 5/12 & 1/12 & 5/12 & 1/12 \\ 5/12 & 1/12 & 5/12 & 1/12 \\ 1/12 & 5/12 & 1/12 & 5/12 \\ 1/12 & 5/12 & 1/12 & 5/12 \end{pmatrix} \qquad \psi = \begin{pmatrix} 6/12 & 0 & 6/12 & 0 \\ 6/12 & 0 & 6/12 & 0 \\ 0 & 6/12 & 0 & 6/12 \\ 0 & 6/12 & 0 & 6/12 \end{pmatrix}.$$

For instance, $\sigma_{17} = 5/12$ is found in the top row, third column. To the right is the corresponding matrix associated to $\Psi = [\mu_1, \mu_2]$, where

$$\mu_1 = \big\{(1,5), (2,7), (3,6), (4,8)\big\} \qquad \mu_2 = \big\{(1,7), (2,5), (3,8), (4,6)\big\}.$$

As Ψ is a Pareto-improvement over Σ, Σ is not sd-efficient.

Appendix C. Tan's (1991) stable partitions

A *partition* of N is $A^1, A^2, \ldots$ such that, for all $i \neq j$, $A^i \cap A^j = \emptyset$ and $\cup_i A^i = N$. As a special case, $\mu \in \mathscr{M}^*$ induces a partition of N into pairs and singletons, $\{1, \mu(1)\}, \{2, \mu(2)\}$, and so on. Tan (1991) considers also larger partition sets. A *ring* is a list of agents $x_1, x_2, \ldots, x_m$ such that, for each x_i, $x_{i+1}\ P_{x_i}\ x_{i-1}$ (mod m). A *stable partition* $A^1, A^2, \ldots$ is such that (i) each partition set A^i is either a single agent, a pair of agents, or corresponds to a ring, and (ii) for each $x_i \in A^k$ and each $y_j \in A^m$ such that $y_j \neq x_{i+1}$,

$$y_j\ P_{x_i}\ x_{i-1} \Rightarrow y_{j-1}\ P_{y_j}\ x_i.$$

If $A^k = \{x_i\}$, then x_{i-1} refers to x_i. If $A^k = \{x_i, x_{i+1}\}$, then x_{i-1} refers to x_{i+1}. Tan (1991) shows that every general pairing problem has a stable partition.