# A Note on Interim Stability in One-to-One Matching Problems

Wouter Vergote[1,a,b]

[a]CORE, Université catholique de Louvain, 34 voie du Roman Pays, B-1348 Louvain-la-Neuve, Belgium.
[b]CEREC, Université Saint-Louis - Bruxelles, Boulevard du Jardin Botanique 43, B-1000 Brussels, Belgium.

07/09/2016

## Abstract

This note introduces interim stability in one-to-one matching problems. Given a status quo matching, two agents may want to form a new partnership without being able to implement it. In this setting we develop two interim stability concepts, direct and (coalition) trade stability, akin to Gale-Shapley stability and exchange stability (Alcalde, 1995) respectively. We show that coalition-trade stability is a refinement of direct stability. When no matching location scarcity exists then direct stability is equivalent to Gale-Shapley stability and coalition-trade stability is equivalent to requiring both exchange stability and Gale-Shapley stability. We partially characterize coalition-trade stable matchings through providing an interesting link between trade dominance and indirect dominance (Harsanyi 1974, Chwe 1994). Building on this, we show that deciding whether the farsighted core of an individually rational roommate problem exists is NP-complete.

**JEL classification:** C71, C78

**Keywords:** One-to-one matching, Interim Stability, Direct Dominance, Trade Dominance, Indirect Dominance.

[1]This note has benefitted from very useful comments from two referees. I gratefully acknowledge financial support from the FNRS, Belgium and from the Belgian French speaking community ARC project #15/20-072 (Social and Economic Network Formation under Limited Farsightedness: Theory and Applications), Université Saint-Louis - Bruxelles, October 2015 - September 2020.

# 1 Introduction

Recently novel solution concepts have been introduced to study stability in one-to-one matching problems depending on the rules that govern how a status quo matching can be altered by a set of agents. In case of the marriage problem, Elhers (2007) studied von Neumann Morgenstern stable sets while Mauleon et al. (2011) characterized farsighted von Neumann Morgenstern stable sets. For the roommate problem, Inarra et al. (2013) analyze absorbing sets while Klaus et al. (2011) study farsighted von Neumann Morgenstern stable sets. These solutions concepts have in common that, given a status quo, deviating agents must be able to effectively implement a matching they wish to deviate to and that when they do so, they face no limits in partnering up. In some settings, this assumption is not realistic, as pointed out by Morrill (2010). Consider two pilots who would rather fly together than flying with their respective partners. If the current assignment provides rights to those partners (to fly on a given route, for instance), then these two pilots may not be able to match up. In sports tournaments rivals need to be assigned to a certain contest venue. In the roommate problem, two students who prefer to share a room rather than sticking to their current roommates can only do so if there is a room available to them. Indeed, when an assignment gives a student the right to reside at a given room, and all other rooms are occupied, these students cannot, on their own, enforce a new matching in which they occupy a room together. Generally, the set of matching locations (routes, rooms, venues, ...) can restrict the possible matchings that agents can enforce over the current assignment. It is then natural to consider the set of available matching locations as a primitive of the matching problem, alongside the agents and their preferences and ask which matchings are interim stable when agents have the opportunity to enforce another matching, respecting the rights the status quo grants to all agents.

The purpose of this note is thus twofold: to provide a novel framework that takes these matching restrictions into account and to derive some insightful connections between this set-up and existing stability concepts in the literature. To do so, we introduce a finite set of matching locations ('bundles of matching rights') and we generalize one-to-one matching problems by requiring that a match between agents must happen at a matching 'location'. We do not have in mind that matchings must physically take place at some location, but rather that the set of locations represents a limit on the amount of possible partnerships. We then require that a match between two agents happens if and only if both are assigned to the same location by a *location mapping*. To focus on the impact of introducing location restrictions in one-to-one matching problems we assume that agents have no preferences over the different locations and, when ending up unmatched, that they are indifferent between being assigned to some location or not. Metaphorically, think of a chess tournament: locations are identical table/chess board combinations. Agents only care about who they are matched against, and when they are not matched to any rival, they are indifferent between sitting at a chess table or sitting on a bench in the park. How interim stability is affected depends on the rules that determine how the status quo can be altered. Two sets of rules are considered. Under the first set of rules (direct enforceability) agents can only trade locations when the other agents assigned to these locations agree. Under the second set of rules (trade enforceability), agents do not need the consent of anyone else to trade their locations. Direct and trade enforceability allow us to define the interim stability concepts of direct stability and (coalition-) trade stability respectively. We then develop natural connections between these stability concepts with previously defined stability concepts in the literature: Gale-Shapley stability, exchange stability, (farsighted) stable set, among others. Gale-Shapley stability, for instance, is an ex ante stability[2] concept in the sense that it requires that, *before* a matching is implemented, no single agent or a pair of agents would like to deviate. This is equivalent to assuming that, under direct stability, the set of matching locations is sufficiently large. A matching is exchange stable (Alcalde, 1995) if there does not exist an exchange blocking pair: no two agents can be made better off by exchanging their current matching position. This is equivalent to (implicitly) assuming that, under trade stability, there are no unassigned matching rights which a set of agents can exchange their current matching

[2]This meaning of ex ante stability is to be distinguished from Kesten and Unver (2015).

rights for. In other words, the set of matching locations must be limited. Our note is rooted in the same spirit as Morrill (2010) who studies the roommate problem and asks: 'ex post, what types of coalitions will be able to block a given assignment?', recognizing that agents face two restrictions: 1. the set of rooms is limited to exactly half the amount of students and 2. bilateral approval is needed to dissolve a match. In Morrill's (2010) set-up Pareto optimal matchings are the candidate stable matchings. Morrill (2010) notes that when the current assignment can be dissolved unilaterally, then (Gale-Shapley) stability is a natural stability concept. Our note argues that direct (trade) stability is a natural solution concept when rights cannot (can) be exchanged. Our first result (proposition 1) is that coalition-trade stability is a refinement of direct stability. We subsequently show (proposition 2) that when matching locations are not scarce then direct stability is equivalent to Gale-Shapley stability and coalition-trade stability is equivalent to requiring Gale-Shapley stability and exchange stability simultaneously. Our next result provides an interesting link between trade dominance and indirect dominance (Harsanyi, 1974, and later formalized by Chwe, 1994). The *farsighted core* of a matching problem is the set of all matchings that are not indirectly dominated by some matching. In Theorem 1 we show that whenever a matching trade dominates some other matching it also indirectly dominates this matching, as long as all agents who see their match change find their new partner acceptable. Intuitively, if two agents wish to exchange their partners but would need the consent of the latter to do so, then they could perform this trade in two steps: in step 1 they leave their current partner and in step 2 they propose to match to complete the trade. If agents are allowed to trade their current partners, these two steps can be done in one. Such indirect dominance path cannot exist if the agents who were forced into a new partnership would rather be single than be matched to their new partner in a trade they did not initiate. In Corollary 1 we show that the set of coalition-trade stable matchings of individually rational matching problems is a superset of the farsighted core and equivalent to the farsighted core if there is no location scarcity. It is well known that in this setting the farsighted core, and hence the set of (coalition) trade stable matchings, is often empty and if it exists, it is a singleton. For this reason the literature (not assuming restrictions on the set of matching locations) introduced alternative stability concepts to study farsightedly stable matchings. A popular stability concept is the *farsighted stable set* (FSS)[3]. A set of matchings is a FSS if it satisfies both internal and external stability with respect to indirect dominance. Equivalently we define a trade stable set (TSS) as a set of matchings that satisfies both internal and external stability with respect to trade dominance. We then find (proposition 3) that a trade stable set can never be a proper subset of a farsighted stable set while it is sometimes a superset of a farsighted stable set. In addition we provide an example of a matching problem without a FSS for which there exists nonetheless a TSS. Lastly, the computer science literature paid ample attention to the *computational complexity* question of determining whether a given matching problem admits a Gale-Shapley stable[4] or exchange stable (Alcalde) matching. It is known (Manlove 2013 and Irving 2008) that deciding whether a (one-to-one) matching is both exchange stable (à la Alcalde 1995) and Gale-Shapley stable is NP complete. The last contribution of this note exploits the link between trade dominance and indirect dominance to say something about the complexity of finding farsightedly stable matchings in one-to-one matching problems when the set of location rights is *sufficiently large*. We find that (corollary 2) while deciding whether a matching is Gale-Shapley stable is solvable in polynomial time, deciding whether a one-to-one matching problem admits a unique farsighted matching (= the farsighted core) is NP-complete. This result bridges recent developments in the computer science literature and economics literature on one-to-one matching problems.

The rest of the note is organized as follows. Section 2 introduces one-to-one matching problems with location restrictions. Section 3 introduces direct stability and (coalition)-trade stability. Section 4 analyzes the relationship between trade dominance and indirect dominance. Section 5 deals with computational complexity questions and Section 6 concludes.

[3]See Ray and Vohra (2015) for a recent discussion of this concept.

[4]Determining whether a given roommate problem with strict preferences admits a stable assignment is solvable in polynomial time (see for instance Irving, 1985 and Ronn, 1990, for a discussion).

# 2 One-to-one matching problems with location restrictions

## 2.1 The primitives

A one-to-one matching problem, or roommate problem, is a triple $(L, N, P)$. $L$ is a finite set of locations, with a specific location denoted by $l \in L$, $N$ is a finite set of agents and $P$ is a preference profile specifying for each agent $i \in N$ a *strict* preference ordering over $N$. That is, $P = \{P(1), ..., P(i), ..., P(n)\}$, where $P(i)$ is agent $i$'s strict preference ordering over the agents in $N$ including herself. For instance, $P(i) = 4, 5, i, 2, ...$ indicates that agent $i$ prefers agent 4 to agent 5 and she prefers to remain alone rather than get matched to anyone else. We denote by $\mathcal{L}$ and $\mathcal{N}$ the cardinality of $L$ and $N$ respectively. We denote by $R$ the weak orders associated with $P$. We write $j \succ_i k$ if agent $i$ strictly prefers $j$ to $k$, $j \sim_i k$ if $i$ is indifferent between $j$ and $k$, and $j \succsim_i k$ if $j \succ_i k$ or $j \sim_i k$. The *primitives* of a one-to-one matching problem are $L, N$ and $P$. A marriage problem with location restrictions is a roommate problem $(L, N, P)$ where $N$ is the union of two disjoint finite sets: a set of men $M = \{m_1, \ldots, m_h\}$, and a set of women, $W = \{w_1, \ldots, w_f\}$, where possibly $h \neq fs$, and $P$ is a preference profile specifying for each man $m \in M$ a strict preference ordering over $W \cup \{m\}$ and for each woman $w \in W$ a strict preference ordering over $M \cup \{w\}$: $P = \{P(m_1), \ldots, P(m_h), P(w_1), \ldots, P(w_f)\}$. That is, each man (woman) prefers being unmatched to be matched with any other agent in $M$ ($W$, respectively). Since the a marriage problem is a special kind of roommate problem we will, throughout the note, use the more general set up and notation of the roommate problem, while sometimes specifically referring to the marriage problem whenever appropriate. A roommate problem is *individually rational* if all agents prefer to be matched to remaining alone: $\forall i, j \in N : j \succ_i i$.

## 2.2 Matching with location restrictions

We now formally introduce the idea that a match between two different agents must happen at a matching 'location'. Define the mapping $\lambda : N \to L \cup \emptyset$ to be a function that assigns a location to each player allowing for the possibility that agents are not assigned to any location $l \in L$:

**Definition 1.** $\lambda : N \to L \cup \emptyset$ is a **location mapping** when $\lambda(i) = \lambda(j) = l \in L$ and $i \neq j \Rightarrow \lambda(k) \neq l$ for all $k \in N \setminus \{i, j\}$

The mapping $\lambda$ assigns a location to at most two agents. It also allows for agents not to be assigned to any location: $\lambda(i) = \emptyset$. Define $\lambda^{-1} : L \to N \cup \emptyset$ as the correspondence that assigns to each location $l$ the set of agents that are located at this location. If no agent is located at $l \in L$, then $\lambda^{-1}(l) = \emptyset$. Let $\lambda \in \Lambda$ where $\Lambda$ is the set of all possible location mappings. The concept of a location mapping allows us to define a matching under location restrictions.

**Definition 2.** Given is $\lambda \in \Lambda$. A **matching** is a function $\mu_\lambda : N \to N$ satisfying the following properties:

1. $\forall i \in N$, $\mu_\lambda(\mu_\lambda(i)) = i$;
2. $\forall\ i \neq j : \mu_\lambda(i) = j \Leftrightarrow \lambda(i) = \lambda(j) = l \in L$.

Condition 1 implies that a matching must yield a partition of the set $N$ into pairs and/or singletons. Condition 2 imposes that for two different agents to be matched they must be assigned to the same location $l \in L$. Denote by $\mathcal{M}^*$ the set of all matchings. One interpretation of a matching problem with location restrictions is that in order to be matched at location $l$, two agents must both possess (be assigned) the matching right attached to matching location $l \in L$. In other words, if an agent 'owns' a matching right to a certain matching location then no other agent can assign herself this matching right unless the 'owner' agrees to this. We assume that agents have no preference over the possible matching locations but only over their possible partners at such a location. We do so to focus on the consequences of introducing scarcity in the set of possible matching locations. Agent $\mu_\lambda(i)$ is agent $i$'s *partner* at $\mu_\lambda$; i.e., the agent with whom she is matched to (possibly herself). A matching $\mu_\lambda$ is *individually rational* if each agent is acceptable to her partner, i.e. $\mu_\lambda(i) \succsim_i i$ for all $i \in N$. A matching problem $(L, N, P)$ is *individually rational* if all $\mu_\lambda \in \mathcal{M}^*$ are individually rational. We extend each agent's preference over her potential partners to the set of matchings in the following way. We say that agent $i$ prefers $\mu'_{\lambda'}$ to $\mu_\lambda$, if and only if agent $i$ prefers her partner at $\mu'_{\lambda'}$ to her partner at $\mu_\lambda$, $\mu'_{\lambda'}(i) \succ_i \mu_\lambda(i)$. Abusing notation, we write this as $\mu'_{\lambda'} \succ_i \mu_\lambda$. A *coalition* $S$ is a subset of $N$. In order to study stability in this setting one needs to know how a given matching $\mu_\lambda$ can be altered by a coalition of agents $S \subset N$. We will introduce two different notions of enforceability: direct enforceability and trade enforceability. A specific (interim) stability concept is then defined using a given enforceability concept.

# 3 Enforceability, dominance and stability

## 3.1 Direct enforceability, dominance and stability

We first develop the concept of direct enforceability according to which agents can reassign locations if all the agents assigned to these locations agree. Given any location assignment $\lambda$, there are some locations that are empty, some that have only one agent and the remaining ones that have two agents assigned to. Direct enforceability implies that members of a coalition $S$ are restricted, in order enforce a new location assignment over the current one, to use only those locations that are not assigned to an 'outsider' of $S$. These locations are thus the empty ones, those that are assigned to only one agent and member of $S$ and those assigned to two members of $S$. Formally define, for any coalition $S$, the set $L_\lambda(S)$ as the locations that members of coalition $S$ can 'directly' control. No one outside $S$ is assigned to a location belonging to $L_\lambda(S)$ :

$$L_\lambda(S) = \left\{ l \in L, \lambda^{-1}(l) \subset S \ \right\}$$

Note that this definition implies that all non-assigned locations are directly controlled by the members of $S$ when contemplating a deviation from the current matching. We can now state the definition of direct enforceability:

**Definition 3.** Given is a matching $\mu_\lambda \in \mathcal{M}^*$. A coalition $S \subseteq N$ is said to **directly enforce** a matching $\mu'_{\lambda'}$ over $\mu_\lambda$, denoted by $\mu_\lambda \to_S \mu'_{\lambda'}$, if the following conditions hold for any agent $i \in N$:

1. $\lambda'(i) \neq \lambda(i) \Rightarrow i \in S$ and $\lambda'(i) \in L_\lambda(S)$.
2. $\lambda'(i) = \lambda(i)$ and $\lambda'(i) = \lambda'(j)$ where $\lambda(i) \neq \lambda(j) \Rightarrow i \in S$ and $\lambda(i) \in L_\lambda(S)$.

Note first that the concept of enforceability does not depend on the preferences of the agents. Condition 1 implies that when an agent obtains a new location assignment, then this agent belongs to coalition $S$ and other members of the coalition $S$ can provide this agent with this new location from the set $L_\lambda(S)$: $\lambda'(i) \in L_\lambda(S)$. Condition 2 implies that when an agent accepts a new partner, while not having changed her own location assignment, then this agent must belong to $S$ and the members of $S$ control the remaining location right at that location or it is currently unassigned: $\lambda(i) \in L_\lambda(S)$. Intuitively, this definition says that in order to 'deviate' to a new matching a set of agents $S$ *cannot reassign* the locations of non-members without their permission. However, they *can reassign* the match of non-members by 'divorcing' from a non-member and taking up a different location, possibly by becoming single. For instance, if $j = \mu_\lambda(i)$ where $i \in S$ and $j \notin S$, then $\lambda'(i) \neq \lambda(i)$ and $\lambda'(j) = \lambda(j)$ implies that $\mu'^{-1}(j) = j$. To illustrate how location scarcity affects direct enforceability consider the following example, adapted from Alcalde (1995):

**Example 1.** *Let* $(L, N, P)$ *where* $L = \{l_1, l_2\}$, $N = \{1, 2, 3, 4\}$ *and* $P(1) = 2, 3, 4; P(2) = 3, 1, 4; P(3) = 1, 2, 4$ *and* $P(4) = 1, 2, 3$*, illustrated as follows:*

| *agent 1* | *agent 2* | *agent 3* | *agent 4* |
|---|---|---|---|
| *2* | *3* | *1* | *1* |
| *3* | *1* | *2* | *2* |
| *4* | *4* | *4* | *3* |

*Consider a matching* $\mu_\lambda$ *with the following location assignment* $\lambda(1) = \lambda(3) = l_1$ *and* $\lambda(2) = \lambda(4) = l_2$ *so that* $\mu_\lambda = (13, 24)$. *This means that agent* 1 *is matched to agent* 3 *at location* $l_1$ *and agent* 2 *is matched to agent* 4 *at location* $l_2$. *Then the set of agents* $S = \{1, 4\}$ *cannot directly enforce another matching over* $\mu_\lambda$ *in which they are matched to one another. If the set of matching locations were* $L = \{l_1, l_2, l_3\}$ *then they would be able to enforce such a matching.*

Direct enforceability allows us to define the concepts of direct dominance and direct stability:

**Definition 4.** Given is a matching problem $(L, N, P)$

1. A matching $\mu_\lambda$ is **directly dominated** by $\mu'_{\lambda'}$ through coalition $S$, denoted by $\mu_\lambda <_S \mu'_{\lambda'}$, if there exists a coalition $S \subseteq N$ of agents such that $\mu'_{\lambda'} \succ_i \mu_\lambda \ \forall i \in S$ and $\mu_\lambda \rightarrow_S \mu'_{\lambda'}$.

2. A matching $\mu_\lambda$ is **directly stable** if no other matching $\mu'_{\lambda'}$ directly dominates $\mu_\lambda$.

A matching $\mu_\lambda$ is *blocked by a coalition* $S \subseteq N$ if there exists a matching $\mu'_{\lambda'}$ such that $\mu'_{\lambda'}(S) = S$ and for all $i \in S$, $\mu'_{\lambda'} \succ_i \mu$. If $S$ blocks $\mu_\lambda$, then $S$ is called a *direct blocking coalition* of $\mu$. If a coalition $S = \{i, j\}$ blocks a matching $\mu_\lambda$, then we call the pair $\{i, j\}$ (possibly $i = j$) a direct blocking pair of $\mu_\lambda$. The *direct core* of a roommate problem with location restrictions, denoted by $C(L, N, P)$, consists of all matchings which are not directly blocked by any coalition.

We now relate this concept to the classic stability concept introduced by Gale and Shapley (1962). According to Gale and Shapley a matching is stable if there does not exist any blocking pair or individual in which a blocking pair is always allowed to deviate by getting matched. When $\mathcal{L} \geqslant \mathcal{N} - 1$, two agents who would prefer to be linked can always find a location that is currently assigned to no one else. It is thus immediate that the concepts of direct stability and Gale-Shapley stability are equivalent whenever $\mathcal{L} \geqslant \mathcal{N} - 1$. When this is the case then location restrictions are essentially immaterial and direct stability yields equivalent predictions as Gale-Shapley stability. When $\mathcal{L} < \mathcal{N} - 1$, agents have fewer opportunities to deviate and the possibility arises that a matching is directly stable while it is not stable in the Gale-Shapley sense. Example 1 illustrates this:

**Example 1 continued.** *Since there exists an odd ring* $\{123\}$[5] *this matching problem is 'unsolvable' in the Gale-Shapley sense. The same conclusion obtains when* $L = \{l_1, ..., l_k\}$ *where* $k > 2$. *However, when* $L = \{l_1, l_2\}$, *the following matchings are directly stable:* $\mu_\lambda = (13, 24), \mu'_{\lambda'} = (14, 23), \mu^*_{\lambda^*} = (12, 34)$. *Imposing more location restrictions leads to more matchings being directly stable. Consider* $(L_1, N, P)$ *where* $L = \{l_1\}$, *then all (constrained*[6]*) Pareto optimal matchings* $\{(12), (13), (14), (23), (24), (34)\}$ *are directly stable.*

## 3.2 Trade enforceability, dominance and stability

We now introduce trade enforceability according to which agents can trade their location rights without the consent of their current partner. It implies that members of a coalition $S$ can, in order enforce a now location assignment over the current one, in addition to the locations in $L_\lambda(S)$, use those locations over which they have only partial control: locations to which only one 'outsider' of $S$ is assigned. The only locations they can not use are those assigned to two agents who do not belong to $S$. Formally define for any coalition $S$ the set $\overline{\mathcal{L}}_\lambda(S)$ as those locations that are assigned

[5] Let $J = j_1, ..., j_k \subseteq N$ be an ordered set of agents. The set $J$ is a ring if $k \geq 3$ and for all $i \in 1, ..., k, j_{i+1} \succ_{j_i} j_{i-1} \succ_{j_i} j_i$ (subscript modulo $k$). A ring $J$ is odd if $k$ is odd.

[6] Constrained in the sense that only one match can possibly be formed.

to two different agents who do not belong to $S$. Then the set of locations coalition $S$ directly *and* partially controls is the complement of $\overline{\mathcal{L}}_\lambda(S)$:

$$\mathcal{L}_\lambda(S) = L \setminus \overline{\mathcal{L}}_\lambda(S)$$

Alternatively, let $\mathcal{L}^1_\lambda(N \setminus S)$ be the set of locations which are assigned to exactly one agent not belonging to $S$. Then equivalently:

$$\mathcal{L}_\lambda(S) = L_\lambda(S) \cup \mathcal{L}^1_\lambda(N \setminus S)$$

$\mathcal{L}_\lambda(S)$ is thus the set of locations that members of $S$ can 'use' to form a new matching through trading their location rights, respecting the location rights of agents who do not belong to $S$. Again, all locations that are not assigned under $\lambda$ belong to this set, but now also those locations that only belong to one agent who does not belong to the set. We now state the definition of trade enforceability[7]:

**Definition 5.** Given is a matching $\mu_\lambda \in \mathcal{M}^*$. A coalition $S \subseteq N$ is said to be able to **trade enforce** a matching $\mu'_{\lambda'}$ over $\mu_\lambda$, denoted by $\mu_\lambda \rightleftharpoons_T \mu'_{\lambda'}$, if for any agent $i \in N$: $\lambda'(i) \neq \lambda(i) \Rightarrow i \in S$ and $\lambda'(i) \in \mathcal{L}_\lambda(S)$.

Intuitively, this definition says that in order to enforce a new matching over the current matching, a set of agents $S$ can do so by reshuffling the available location rights among members of $S$. This can happen in two ways. First, an agent can obtain a new allocation right from another agent of the set $S$ through a permutation of location rights among members of $S$. Second, the agent can obtain a location right that was not assigned to anyone under $\lambda$ or simply give up their current assignment without being assigned a new one according to $\lambda'$. The members of the set S cannot reassign the locations of non-members, but they can reassign the partner of non-members by exchanging their location right with some other agent or by simply exchanging their location for a currently available location[8]. Trade enforceability implies that if a partner x of an agent y exchanges her matching right with someone else, say agent z, then agent y has no say in this exchange, even if it means that she will become matched to agent z. The key difference between the two enforceability conditions is that direct enforceability requires that the agents must control all rights assigned to a location, while trade enforceability only requires that they exchange location rights, thereby potentially 'forcing' other agents to accept a different partner from the (deviating) set. It follows that direct enforceability puts more restrictions on what a set of agents $S$ can do in order to change a matching.

**Lemma 1.** *Suppose $\mu_\lambda \rightarrow_S \mu'_{\lambda'}$ then there exists $S' \subseteq S$ such that $\mu_\lambda \rightleftharpoons_S \mu'_{\lambda'}$.*

*Proof.* Note that $\mu_\lambda \rightarrow_S \mu'_{\lambda'}$ implies that for all $i$ such that $\lambda'(i) \neq \lambda(i) \Rightarrow i \in S$ and $\lambda'(i) \in L_\lambda(S)$. Since $L_\lambda(S) \subseteq \mathcal{L}_\lambda(S)$, let all these $i$ belong to $S'$. But then, by the definition of trade enforceability: $\mu_\lambda \rightleftharpoons_{S'} \mu'_{\lambda'}$ where we note that $S' = S$. $\square$

Trade enforceability allows us to define the concept of trade dominance:

**Definition 6.** A matching $\mu_\lambda$ is trade dominated by $\mu'_{\lambda'}$ by coalition $S$, denoted by $\mu_\lambda \lhd_S \mu'_{\lambda'}$, if $\mu_\lambda \rightleftharpoons_S \mu'_{\lambda'}$ and $\mu'_{\lambda'} \succ_i \mu_\lambda\ \forall i \in S$.

When a coalition trade dominates some matching we say that it **trade blocks** this matching. When a trade blocking coalition is a pair (singleton) it is called a **trade blocking pair** (singleton).

**Definition 7.** A matching $\mu_\lambda$ is **trade stable** if it is not trade dominated by any pair or individual. A matching $\mu_\lambda$ is **coalition-trade stable** if it is not trade dominated by any coalition S.

[7]Note again that the notion of trade enforceability does not depend on the preferences of the agents.

[8]Note that we allow agents to exchange their current location right for any available (unassigned) location right as it seems natural to let agents perform this trade when they are allowed to exchange matching location rights.

The definition of trade blocking is broader than the notion of exchange blocking introduced by Alcalde (1995): a pair of agents $\{i,j\}$ is said to *exchange block* a matching when, only by exchanging their current match, they both improve upon their current match. Following Alcalde (1995), a matching is (coalition) exchange stable if there does not exist an exchange blocking pair or coalition:

**Definition 8.** (Alcalde, 1995) A matching $\mu_\lambda \in \mathcal{M}^*$ is *exchange stable* if $\nexists\{i,j\}$ where $\mu_\lambda(i) \neq i, \mu_\lambda(j) \neq j \mu'_{\lambda'} \succ_i \mu_\lambda \succ_i \mu_\lambda$ such that $\mu_\lambda(j) \succ_i \mu_\lambda(i)$ and $\mu_\lambda(i) \succ_j \mu_\lambda(j)$. Matching $\mu_\lambda \in \mathcal{M}^*$ is *coalition-exchange stable* if $\nexists S \subseteq N$ such that $\forall i \in S : \mu_\lambda(i) \neq i$ and $\forall i \in S : \exists j \in S : \mu_\lambda(j) \succ_i \mu_\lambda(i)$.

Alcalde (1995) assumes that only 'partner trades' can happen and does not, implicitly, consider 1) that agents may decide to become single by giving up their current matching right and 2) that there could be available matching locations that agents can use to become matched by exchanging their current location right for such an available location right. The concept of trade enforceability takes the latter two possibilities into account. We now illustrate these concepts in example 1 again:

**Example 1 continued** *In the example, only the matching* $\{(13,24)\}$ *is exchange stable according to Alcalde (1995). We obtain that*

1. *when* $L = \{l_1\}$*, then all individually rational matchings* $\{(12),(13),(14),(23),(24),(34)\}$ *are trade stable. This set is equivalent to the set of directly stable matchings.*
2. *when* $L = \{l_1, l_2\}$*, then only matching* $\{(13,24)\}$ *is trade stable.*
3. *when* $L = \{l_1, l_2, ..., l_K\}$ *where* $K > 2$*, then no matching is trade stable.*

The computer science literature introduced a different stability notion which simultaneously requires exchange stability and Gale-Shapley stability[9]: a matching is both (coalition-) exchange stable and Gale-Shapley stable if there is no exchange blocking pair (coalition) and no blocking pair. Requiring both Gale-Shapley stability, implicitly assuming no location scarcity, and exchange stability, implicitly assuming location scarcity, seems a somewhat peculiar assumption. We propose to consider the set of property rights (available locations) as a primitive of the model and let the stability concept be based on the enforceability rules. In particular, can matching rights be obtained through an exchange of location rights or not? When they can (not), we argue that the concept of trade (direct) enforceability is appropriate.

### 3.3 Direct and coalition trade stability

Example 1 is illustrative of the property that coalition-trade stability is a refinement of direct stability. It is insightful to contrast this to the conclusion made by Alcalde (1995) that Gale-Shapley stability and exchange stability are mutually independent concepts. Alcalde's conclusion is based, implicitly, on the assumption that the set of available matching rights is different when testing exchange stability and Gale-Shapley stability. Once we fix the set of matching locations, then coalition-trade stability implies direct stability. We now show formally that trade stability is a refinement of direct stability. We first show the following lemma:

**Lemma 2.** *When a matching* $\mu_\lambda$ *is directly blocked by a* $\{i,j\}$ *then it is also trade blocked by* $\{i,j\}$*.*

*Proof.* First assume that $i = j$. Hence for agent $i$ we have that $i \succ \mu_\lambda(i)$. Then agent $i$ can simply 'divorce' from $\mu_\lambda(i)$ without being assigned a new location. But the same move can be done by a trade blocking singleton in which agent $i$ just gives up her location assignment. Second assume that $i \neq j$ and $\mu'_{\lambda'}(i) = j$. Then it must be that $\lambda'(i) = \lambda'(j) = l'$ where $l' \in L_\lambda(\{i,j\})$. In other words, location $\lambda^{-1}(l') = \emptyset$ was not assigned to anyone in $\lambda$. But then $\{i,j\}$ can trade enforce $\mu'_{\lambda'}$ over $\mu_\lambda$, by exchanging their current location for $l'$. □

[9] See Chechlarova and Manlove (2005), Irving (2008), McDermid et al. (2007), Manlove (2013).

Lemma 2 can be generalized: direct dominance implies trade dominance.

**Proposition 1.** *Given a one-to-one matching problem (N,L,P). Then* $\mu_\lambda <_S \mu'_{\lambda'} \Rightarrow \exists S' : \mu_\lambda \lhd_{S'} \mu'_{\lambda'}$.

*Proof.* Let $\bar{S} = \{i \in S : \lambda(i) = \lambda'(i)\ and\ \exists j \in S\ such\ that\ \lambda'(j) = \lambda(i) \neq \lambda(j)\}$. The set $\bar{S}$ consists of agents who remain at their original position and obtain a new partner from the set $S$ in $\mu'_{\lambda'}$. Then consider $S' = S \setminus \bar{S} \subseteq S$. Then all locations which have seen a permutation of agents in going from $\mu_\lambda$ to $\mu'_{\lambda'}$ are directly and partially controlled by members of $S'$ such that members of $S'$ can exchange enforce $\mu'_{\lambda'}$ over $\mu_\lambda$, and all be better off: $\mu_\lambda \lhd_{S'} \mu'_{\lambda'}$. □

We return briefly to the relationship between (coalition-) trade stability and other stability concepts. We obtain that whenever there is no scarcity in locations ($\mathcal{L} \geq \mathcal{N}-1$) requiring the absence of trade blocking pairs is equivalent to *simultaneously* requiring the absence of exchange blocking pairs[10] (exchange stability à la Alcalde, 1995) and the absence of blocking pairs (stability à la Gale-Shapley, 1962). As mentioned above, requiring both Gale-Shapley and exchange stability is a stability concept used in the computer science literature.

**Proposition 2.** *Given is a one-to-one matching problem (L,N,P). When* $\mathcal{L} \geq \mathcal{N}-1$, *then (coalition) trade stability is equivalent to requiring both Gale-Shapley stability and (coalition) exchange stability (Alcalde, 1995).*

*Proof.* The proof is done for trade stability and exchange stability. Proving that it also holds for coalition-trade stability and coalition exchange stability follows the exact same lines and is omitted.

$\Rightarrow$ Suppose that $\mu_\lambda$ is trade stable and there exists either a blocking pair (individual) or an exchange blocking pair. Suppose first that $\{i, j\}$ is a blocking pair (or individual when $i = j$) of $\mu_\lambda$. Since $\mathcal{L} \geq \mathcal{N}-1$ , then $\{i, j\}$ can enforce the matching $\mu'_{\lambda'}$ where $\mu'_{\lambda'} = \mu_\lambda - i\mu_\lambda(i) - j\mu_\lambda(j) + ij$. We then have that $\mu_\lambda \lhd_{\{i,j\}} \mu'_{\lambda'}$, a contradiction. Now suppose that $\{i, j\}$ is an exchange blocking pair. But then for any $\mathcal{L}$ we have that $\mu_\lambda \lhd_{\{i,j\}} \mu'_{\lambda'}$, again a contradiction.

$\Leftarrow$ Suppose there does not exist a blocking pair, nor an exchange blocking pair but their exists a pair $\{i, j\}$ (or individual when $i = j$) and a matching $\mu'_{\lambda'}$ such that $\mu_\lambda \lhd_{\{i,j\}} \mu'_{\lambda'}$. Since $\{i, j\}$ is not an exchange blocking pair or individual, then it must be that they are matched or single in $\mu'_{\lambda'}$ in which they are better off. But then $\{i, j\}$ would be a blocking pair, a contradiction □

# 4 Characterizing coalition-trade stable matchings

## 4.1 A link between coalition trade dominance and indirect dominance

We can now be more precise about the matchings that are (coalition) trade stable. Consider the following example:

**Example 2.** *Let* $(L, N, P)$ *where* $L = \{l_1, l_2\}$ , $N = \{1, 2, 3, 4\}$ *and* $P(1) = 3, 4, 1; P(2) = 4, 3, 2; P(3) = 2, 1, 3$ *and* $P(4) = 1, 2, 4$. *Given these preferences this is equivalent to a marriage problem. These preferences are illustrated as follows:*

| *agent 1* | *agent 2* | *agent 3* | *agent 4* |
|---|---|---|---|
| *3* | *4* | *2* | *1* |
| *4* | *3* | *1* | *2* |
| *1* | *2* | *3* | *4* |

*Consider matchings* $\mu_\lambda = (13, 24)$ *and* $\mu'_{\lambda'} = (14, 23)$. *These two matchings are directly stable and stable in the Gale-Shapley sense. However, these matchings are not trade stable nor exchange stable in the sense of Alcalde (1995). Indeed, agents 3 and 4 can trade enforce matching* $\mu'_{\lambda'}$ *over matching* $\mu_\lambda$ *and agents 1 and 2 can trade enforce matching* $\mu_\lambda$ *over matching* $\mu'_{\lambda'}$.

[10] Note that coalition-trade stability is a refinement of trade stability.

By close inspection we observe that while agents *1* and *2* cannot directly enforce matching $\mu_\lambda$ over matching $\mu'_{\lambda'}$, they may do so in two steps, assuming that both agents are forward-looking. In a first step they could divorce in order to, in a second step, match with each others ex-partners. Hence, if agents *1* and *2* cannot exchange their matching location rights, they may still be able to 'trade' their partners if they are not myopic. We generalize this intuition by showing that there is a close relationship between trade dominance and indirect dominance, a dominance concept introduced by Harsanyi (1974) and Chwe (1994) in order to study what happens when agents do not care about the immediate consequences of their actions but rather about the final outcome after other agents have reacted to their initial reaction. A matching $\mu'_{\lambda'}$ *indirectly dominates* $\mu_\lambda$ if $\mu'_{\lambda'}$ can replace $\mu_\lambda$ in a sequence of matchings, such that at each matching along the sequence all deviators are strictly better off at the end matching $\mu'_{\lambda'}$ compared to the status-quo they face. Formally, indirect dominance[11] is defined as follows.

**Definition 9.** A matching $\mu_\lambda$ is indirectly dominated by $\mu'_{\lambda'}$, denoted by $\mu_\lambda \ll \mu'_{\lambda'}$, if there exists a sequence of matchings $\mu^0_{\lambda^0}, \mu^1_{\lambda^1}, ..., \mu^K_{\lambda^K}$ (where $\mu^0_{\lambda^0} = \mu_\lambda$ and $\mu^K_{\lambda^K} = \mu'_{\lambda'}$) and a sequence of coalitions $S^0, S^1, ..., S^{K-1}$ such that for any $k \in \{1, ..., K\}$,

1. $\mu^K_{\lambda^K} \succ_i \mu^{k-1}_{\lambda^{k-1}}$ $\forall i \in S^{k-1}$; and
2. coalition $S^{k-1}$ can enforce the matching $\mu^k_{\lambda^k}$ over $\mu^{k-1}_{\lambda^{k-1}}$.

Direct dominance can be obtained directly from Definition 9 by setting $K = 1$. Obviously, if $\mu_\lambda < \mu'_{\lambda'}$ then $\mu_\lambda \ll \mu'_{\lambda'}$; i.e., direct dominance implies indirect dominance. The set of matchings that are not indirectly dominated by any other matching is the farsighted core:

**Definition 10.** A matching $\mu_\lambda$ belongs to the **farsighted core** ('FC') if no other matching $\mu'_{\lambda'}$ indirectly dominates $\mu_\lambda$.

Indirect dominance offers the possibility to agents to contemplate an 'exchange' of their partners, as long as the latter would prefer to remain matched compared to being single. Trade dominance allows these agents to do so directly, even if their partners would prefer to be single rather than being rematched through a trade that they did not initiate. We now confirm this intuition in Theorem 1: whenever an individually rational matching trade dominates some other matching then it also indirectly dominates this matching.

**Theorem 1.** *Let $(L, N, P)$ be a one-to-one matching problem with location restrictions. Let $\mu'_{\lambda'}, \mu_\lambda \in \mathcal{M}^*$, if $\mu'_{\lambda'} \rhd \mu_\lambda$ and if for all $i$ such that $\mu'_{\lambda'}(i) \neq \mu_\lambda(i)$ it is the case that $\mu'_{\lambda'} \succcurlyeq_i i$, then $\mu'' \gg \mu_\lambda$.*

*Proof.* Let $B(\mu_\lambda, \mu'_{\lambda'})$ be the set of agents who are better off in $\mu'_{\lambda'}$ compared to $\mu_\lambda$:

$$B(\mu_\lambda, \mu'_{\lambda'}) = \{i \in N, \mu_\lambda(i) \prec_i \mu_{\lambda'}(i)\}.$$

Let $I(\mu_\lambda)$ be the set of agents who are single in $\mu_\lambda$:

$$I(\mu_\lambda) = \{i \in N, \mu_\lambda(i) = i\}.$$

Then there exists a set of agents $S$ who can trade enforce $\mu'_{\lambda'}$ over $\mu_\lambda$ and be better off in $\mu'_{\lambda'}$. We now construct an indirect dominance path from $\mu_\lambda$ to $\mu'_{\lambda'}$. Let $\mu^1_{\lambda^1}$ be a matching where all agents of $S$ are single, if necessary by leaving their partner in $\mu_\lambda$ by giving up their location assignment under $\lambda$. Let $S_1 \subseteq S$ be those agents who belong to $S$ and have a partner in $\mu_\lambda$ : $S_1 = S \setminus I(\mu_\lambda)$. We then have that $\mu_\lambda \rightarrow_{S_1} \mu^1_{\lambda^1}$ and for all $i \in S_1$ :$\mu'_{\lambda'} \succ_i \mu$. Now consider the set $S_2 = B(\mu^1_{\lambda^1}, \mu'_{\lambda'})$. For any $i \in S_2$ we have:

[11] Note that we do not impose further restrictions on what agents are allowed to do after altering the status quo. In particular, one could imagine that when an agent chooses to give up its location assignment (e.g. a student leaving the campus housing system) she cannot enter the system again. Such additional restrictions would alter the set of possible dynamic deviations.

1. $\mu'_{\lambda'}(i) \neq i$ and $\mu'_{\lambda'}(i) = j \in S$. Then it must be that $\lambda'(i) = \lambda'(j) = l' \in \mathcal{L}_\lambda(\{i, j\})$. But then $\lambda^{-1}(l') = \{i, j\}$ hence $l' \in \mathcal{L}_{\lambda^1}(\{i, j\})$; or

2. $\mu'_{\lambda'}(i) \neq i$ and $\mu'_{\lambda'}(i) = j \notin S$. Then it must be that $\lambda'(i) = \lambda'(j) = l \in \mathcal{L}_\lambda(S)$. But then $\lambda^{-1}(l') = \{i, j\}$ hence $l' \in \mathcal{L}_{\lambda^1}(\{i, j\})$; or

3. $\mu'_{\lambda'}(i) = i$. But then $i \notin S_2$.

We then have that for all $i \in S_2 : \lambda'(i) \in L_{\lambda^1}(S_2)$ and hence we have that $\mu^1_{\lambda^1} \rightarrow_{S_2} \mu'_{\lambda'}$. We conclude that $\mu'_{\lambda'} \gg \mu_\lambda$. $\square$

Theorem 1 implies that for individually rational matching problems the farsighted core must belong to the set of coalition-trade stable matchings. Example 3 shows that this result does not carry over to the case of matchings problems which are not individually rational: when $\mu'_{\lambda'} \rhd \mu_\lambda$ where $\mu'_{\lambda'}$ is not individually rational, it is not necessarily the case that $\mu'_{\lambda'} \gg \mu_\lambda$.

**Example 3.** *Consider the marriage problem $(L, M, W, P)$ where $M = \{m_1, m_2\}$ and $W = \{w_1, w_2\}$ and $L = \{l_1, l_2\}$ with the following preferences:*

| $m_1$ | $m_2$ | $w_1$ | $w_2$ |
|---|---|---|---|
| $w_2$ | $w_1$ | $m_1$ | $m_2$ |
| $w_1$ | $w_2$ | $w_1$ | $w_2$ |
| $m_1$ | $m_2$ | $m_2$ | $m_1$ |

*Let $\mu'_{\lambda'} = (m_1 w_2, m_2 w_1)$, $\mu_\lambda = (m_1 w_1, m_2 w_2)$. We then have that $\mu_\lambda \lhd \mu'_{\lambda'}$ but not that $\mu_\lambda \ll \mu'_{\lambda'}$. Indeed, the women would never accept to remarry a different man than their partner in $\mu_\lambda$. That is, trade enforceability can transform an individually rational match into an individually irrational match. This possibility is ruled out by indirect dominance. Note that in this example the farsighted core is a singleton: $FC = \{\mu_\lambda\}$ while the set of coalition-trade stable matchings is empty.*

While Theorem 1 implies that for individually rational matching problems trade dominance entails indirect dominance, the converse is not the case: indirect dominance does not imply trade dominance. This is illustrated by example 4:

**Example 4.** *Consider the marriage problem $(L, M, W, P)$ where $M = \{m_1, m_2\}$ and $W = \{w_1, w_2\}$ and $L = \{l_1, l_2\}$ with the following preferences:*

| $m_1$ | $m_2$ | $w_1$ | $w_2$ |
|---|---|---|---|
| $w_2$ | $w_2$ | $m_1$ | $m_1$ |
| $w_1$ | $w_1$ | $m_2$ | $m_2$ |
| $m_1$ | $m_2$ | $w_1$ | $w_2$ |

*Let $\mu'_{\lambda'} = (m_1 w_2, m_2 w_1)$ and $\mu_\lambda = (m_1 w_1, m_2 w_2)$. We then have that $\mu_\lambda \ll \mu'_{\lambda'}$ but not that $\mu_\lambda \lhd \mu'_{\lambda'}$ since $m_1$ and $w_2$ cannot obtain the rights to a matching location at which they can match. In this example the farsighted core is a singleton: $FC = \{\mu_\lambda\}$ while the set of trade stable matchings is a couple: $\{\mu_\lambda, \mu'_{\lambda'}\}$.*

Example 4 clarifies that for individually rational matching problems the farsighted core can be a strict subset of the set of trade stable matchings. We now show that this result depends on the level of location scarcity.

**Definition 11.** Given is matching problem $(L, N, P)$. A matching $\mu_\lambda$ belongs to the *set of trade stable matchings* $(T)$ if no pair (or individual) of agents can enforce a matching that trade dominates $\mu_\lambda$ . A matching $\mu_\lambda$ belongs to the *set of coalition-trade stable matchings* $(T^*)$ if no other matching $\mu'_{\lambda'}$ trade dominates $\mu_\lambda$ .

When there is no scarcity ($\mathcal{L} \geq \mathcal{N}{-}1$) the set of coalition-trade stable matchings is equivalent to the farsighted core, while not necessarily equal to the set of trade stable matchings. We have the following corollary:

**Corollary 1.** *The farsighted core of any individually rational matching problem belongs to the set of trade stable matchings: $FC \subset T$. However, $T \subsetneq FC$. When $\mathcal{L} \geq \mathcal{N}{-}1$ we have that $FC = T^* \subsetneq T$.*

*Proof.* Given is that $(L, N, P)$ is individually rational.

1. Suppose first that $\mu_\lambda \in FC$ and $\mu_\lambda \notin T$. Then there exists $\mu'_{\lambda'}$ and $i, j$ where possibly $i = j$ such that $\mu_\lambda \triangleleft_{\{i,j\}} \mu'_{\lambda'}$ . Since $\mu'_{\lambda'}$ is individually rational, it follows from Theorem 1 that $\mu'_{\lambda'} \gg \mu_\lambda$, a contradiction. That $T \subsetneq FC$ follows from example 4.
2. Now let $\mathcal{L} \geq \mathcal{N}{-}1$ , and let $\mu_\lambda \in T^*$. Suppose that there exists $\mu'_{\lambda'}$ such that $\mu'_{\lambda'} \gg \mu_\lambda$. Consider the set $B(\mu_\lambda, \mu'_{\lambda'})$, then $\mu_\lambda \ntriangleleft_{B(\mu_\lambda,\mu'_{\lambda'})} \mu'_{\lambda'}$. However, since $\mathcal{L} \geq \mathcal{N}{-}1$ , there are always enough matching locations to let the members of $B(\mu_\lambda, \mu'_{\lambda'})$ enforce any partner trade and/or any direct blocking coalition since there are at least $\frac{1}{2}\sharp B(\mu_\lambda, \mu'_{\lambda'})$ locations available to members of $B(\mu_\lambda, \mu'_{\lambda'})$ who want to be matched to each other. We have that $\mu_\lambda \triangleleft_{B(\mu_\lambda,\mu'_{\lambda'})} \mu'_{\lambda'}$, a contradiction. Example 4 illustrates that $FC = T^* \nsubseteq T$ even if $\mathcal{L} \geq \mathcal{N}{-}1$.

□

## 4.2 Stable Sets

Often times the farsighted core is empty which lead people to introduce alternative stability concepts to study farsightedly stable matchings. A popular stability concept is that of the *farsighted stable set* (FSS)[12]. A farsighted stable set of a matching problem is a set of matchings that satisfies **internal stability** - no matching of the set indirectly dominates another matching of the set - and **external stability** - all matchings outside the set are indirectly dominated by some matching(s) belonging to the set.

**Definition 12.** A set of matchings $V \subseteq \mathcal{M}^*$ is a *von Neumann Morgenstern farsighted stable set* (FSS) if

1. for all $\mu_\lambda \in V$ , there does not exist $\mu'_{\lambda'} \in V$ such that $\mu'_{\lambda'} \gg \mu_\lambda$ (internal stability);
2. for all $\mu'_{\lambda'} \notin V$ there exists $\mu_\lambda \in V$ such that $\mu_\lambda \gg \mu'_{\lambda'}$ (external stability).

In general, existence of a such a set is not guaranteed, nor its uniqueness when it exists. In the case of no scarcity, Mauleon et al. (2011) and Klaus et al. (2011) have shown that if a matching is Gale-Shapley stable (and thus directly stable in our setting), then it is a singleton FSS. When a matching problem is unsolvable (no Gale-Shapley stable matching exists), a FSS may not exist, as illustrated by our example 1, or may have more than two elements (see example 2 in Klaus et al., 2011).

**Example 1 Continued:** *It is easily verified that no FSS exists.*

We define a *Trade Stable Set* (TSS) as a set of matchings such that they do not trade dominate each other while any matching outside the set is trade dominated by some matching in the set.

**Definition 13.** A set of matchings $\mathit{\Gamma} \subseteq \mathcal{M}^*$ is a *Trade Stable Set* (TSS) if

1. for all $\mu_\lambda \in \mathit{\Gamma}$ , there does not exist $\mu'_{\lambda'} \in \mathit{\Gamma}$ such that $\mu'_{\lambda'} \triangleright \mu_\lambda$ (internal stability);
2. for all $\mu'_{\lambda'} \notin \mathit{\Gamma}$ there exists $\mu_\lambda \in \mathit{\Gamma}$ such that $\mu_\lambda \triangleright \mu'_{\lambda'}$ (external stability).

[12]See Ray and Vohra (2015) for a recent analysis of the concept of farsighted stable set in coalition formation problems.

When $\mathcal{L} \geq \mathcal{N}-1$, Klaus et al. (2011)[13] have shown that any matching belonging to a FSS must be individually rational. Their result immediately extends to our setting with location restrictions.

**Lemma 3.** *Given is a matching problem $(L, N, P)$. Let $V$ be a FSS, then any $\mu_\lambda \in V$ is individually rational.*

*Proof.* The proof is equivalent to the proof of lemma 1 in Klaus et al. (2011) and is therefore omitted. □

We now show (proposition 3) that a TSS cannot be a strict subset of a FSS while the opposite can hold.

**Proposition 3.** *Given is a matching problem $(L, N, P)$. Let $V$ be a FSS and consider $V' \subsetneq V$, then $V'$ is not a TSS.*

*Proof.* Let $V' \subset V$ where $V$ is a FSS. Let $\mu_\lambda \in V$ while $\mu_\lambda \notin V'$. Then there exists $\mu'_{\lambda'} \in V'$ such that $\mu'_{\lambda'} \rhd \mu_\lambda$ , but since $\mu'_{\lambda'}$ is individually rational (using Lemma 3) we know (using Theorem 1) that $\mu'_{\lambda'} \gg \mu_\lambda$, violating internal stability of $V$, a contradiction. □

These results are illustrated below by means of example 4.

**Example 4 Continued:** *Consider the marriage problem $(L, M, W, P)$ where $M = \{m_1, m_2\}$ and $W = \{w_1, w_2\}$ and $L = \{l_1, ..., l_k\}$ where $k > 0$, with the following preferences:*

| $m_1$ | $m_2$ | $w_1$ | $w_2$ |
|---|---|---|---|
| $w_2$ | $w_2$ | $m_1$ | $m_1$ |
| $w_1$ | $w_1$ | $m_2$ | $m_2$ |
| $m_1$ | $m_2$ | $w_1$ | $w_2$ |

*For all sets of location restrictions there is a unique FSS and a unique TSS. For all $k > 1$ the FSS is a strict subset of the TSS.*

1. *When $L = l_1$ we have that $FSS = TSS = \{(m_1w_1), (m_1w_2), (m_2w_1), (m_2w_2)\}$;*
2. *When $L = \{l_1, ..., l_k\}$ where $k > 1$, have that $FSS \subsetneq TSS$ since $FSS = \{(m_1w_2, m_2w_1)\} \subsetneq \{(m_1w_2, m_2w_1), (m_1w_1, m_2w_2)\} = TSS$.*

We end our discussion of the trade stable set by providing an example of a matching problem which has no FSS but does have a TSS.

**Example 1 Continued:** *As mentioned above, this matching problem does not have a FSS. Now consider the following matchings: $\mu^1_{\lambda^1} = (13, 24), \mu^2_{\lambda^2} = (12, 34)$ and $\mu^3_{\lambda^3} = (14, 23)$, in which agents are matched at locations $l_1$ and $l_2$ (all other locations are available), and let $Q$ be the set of all location permutations of these matchings (matched at $l_1$ and $l_2$). There are $\begin{pmatrix} k \\ 2 \end{pmatrix}$ matchings in $Q$ such that no matching of $Q$ trade dominates another matching of $Q$. Equally, all other matchings are trade dominated by a matching of $Q$. Hence, $Q$ is a trade stable set.*

# 5 Computational complexity

In computer science the algorithmic aspects of matching problems have been studied at length. A large body of literature[14] emerged studying whether a given roommate problem admits a (Gale-Shapley) stable matching. A smaller[15] literature asks the same question while replacing Gale-Shapley stability with the concept of exchange stability introduced by Alcalde (1995). Irving

[13]See lemma 1 in Klaus et al., 2011

[14]See Gusfield and Irving, 1989 and Manlove, 2013 for excellent surveys.

[15]Cechlarova (2002), Cechlarova and Manlove (2005), Irving (2008), McDermid et al. (2007)

(2008) finds that deciding whether a roommate problem admits a Gale-Shapley stable matching that is also exchange stable à la Alcalde (1995) is computationally hard:

**Theorem 2.** *(Irving, 2008) The problem of deciding whether a given stable roommate instance admits a stable matching that is exchange stable is NP-complete.*

Our proposition 2 shows that, when $\mathcal{L} \geq \mathcal{N}-1$, coalition-trade stability is equivalent to simultaneously requiring Gale-Shapley stability and coalition-exchange stability. Our corollary 1 shows that the set of coalition-trade stable matchings is equivalent to the farsighted core for an individually rational matching problem. From this we conclude that finding a farsightedly stable matching in an individually rational one-to-one matching problem without location scarcity is also computationally hard:

**Corollary 2.** *Let $\mathcal{L} \geq \mathcal{N}-1$. Deciding whether an individually rational roommate problem admits a farsightedly stable matching is NP complete.*

*Proof.* This follows immediately from Proposition 2, Corollary 1 and Theorem 2 in Irving (2008). □

# 6 Conclusion

This note contributes to the literature on one-to-one matching problems by introducing matching restrictions which form the basis of two interim-stable solution concepts - direct and (coalition)-trade stability - and by establishing insightful links with other stability concepts and the computational complexity literature on matching problems. Many questions remain unanswered. We have not fully characterized trade stable matchings. We have not tackled the question whether a trade stable set always exists. While we have shown that deciding whether the (unique) farsighted core of a individually rational roommate problem exists is computationally hard, we have not done so for individually irrational roommate problems. Additionally, our novel set-up can be extended in many directions. What if locations have multiple location rights? What if agents can only have access to specific locations (e.g. eligibility criteria based on income, age, health status, ...)? How about many-to-one or many-to-many matching problems? We leave these questions and extensions for future research.

# References


**Alcalde** J. (1995) Exchange-proofness or divorce-proofness? Stability in one-sided matching markets, Economic Design 1: 275-287.

**Cechlarova** K. (2002) On the complexity of exchange stable roommates, Discrete Applied Mathematics, 116: 279-287.

**Cechlarova** K. and D.F. Manlove (2005) The exchange-stable marriage problem, Discrete Applied Mathematics, 152: 109-122.

**Chwe** M.S. (1994) Farsighted coalitional stability, Journal of Economic Theory 63: 299–325.

**Diamantoudi** E. and L. Xue (2003) Farsighted stability in hedonic games, Social Choice and Welfare 21: 39-61.

**Ehlers** L. (2007) Von Neumann-Morgenstern stable sets in matching problems. Journal of Economic Theory, 134, 537-547.

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

**Gusfield** D. and R.W. Irving (1989) The stable marriage problem: structure and algorithms, MIT Press.

**Harsanyi** J.C. (1974) An equilibrium-point interpretation of stable sets and a proposed alternative definition, Management Science 20: 1472-1495.

**Irving** R.W. (1985) An efficient algorithm for the 'stable roommates' problem, Journal of Algorithms 6, 577-595. 19

**Irving** R.W. (2008) Stable matching problems with exchange restrictions, Journal of Combinatorial Optimization, Volume 16, Number 4: 344-360.

**Inarra** E., Larrea C. and Molis, E. (2013) Absorbing sets in roommate problems, Games and Economic Behavior 81, 165-178.

**Kesten** O. and M. U¨nver (2015) A theory of school choice lotteries, Theoretical Economics: 543–595

**Klaus** B., Klijn F. and M. Walzl (2011) Farsighted stability for roommate markets., Journal of Public Economic Theory 13: 921-933.

**Manlove** D.F. (2013) Algorithmics of Matching under Preferences, Series on Theoretical Computer Science, Vol. 2, World Scientific.

**Mauleon** A., Vannetelbosch V. and W. Vergote (2011) von Neumann Morgernstern farsightedly stable sets in two-sided matching, Theoretical Economics 6: 499-521.

**McDermid** E., Cheng, C. and I. Suzuki (2007) Hardness results on the man-exchange stable marriage problem with short preference lists, Information procSSSing letters, 101: 13-19.

**Morril** T. (2010) The roommates problem revisited, Journal of Economic Theory, 145, 1739-1756.

**Ray** D. and R. Vohra (2015) The Farsighted Stable Set, Econometrica, Econometric Society, vol. 83(3), pages 977-1011.

**Ronn** E. (1990) NP-complete stable matching problems, journal of algorithms, 11, 285-304.