Stable sets in matching problems with coalitional sovereignty and path dominance

P. Jean-Jacques Herings [a,*], Ana Mauleon [b,c], Vincent Vannetelbosch [c]

[a] Department of Economics, Maastricht University, Maastricht, The Netherlands
[b] CEREC, Saint-Louis University, Brussels, Belgium
[c] CORE, University of Louvain, Louvain-la-Neuve, Belgium

ARTICLE INFO

ABSTRACT

We study von Neumann Morgenstern stable sets for one-to-one matching problems under the assumption of coalitional sovereignty (C), meaning that a deviating coalition of players does not have the power to arrange the matches of agents outside the coalition. We study both the case of pairwise and coalitional deviations. We argue further that dominance has to be replaced by path dominance (P) along the lines of van Deemen (1991) and Page and Wooders (2009). This results in the pairwise CP vNM set in the case of pairwise deviations and the CP vNM set in the case of coalitional deviations. We obtain a unique prediction for both types of stable sets: the set of matchings that belong to the core.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

In the one-to-one matching model known as the marriage problem, there are two disjoint sets of agents, say men and women. The problem is to match agents from one side of the market with agents from the other side, whereas each agent also has the possibility of remaining single. We refer to Roth and Sotomayor (1990) for a comprehensive overview on two-sided matching problems.

For marriage problems, stability is considered to be a central property. A matching is stable if each agent on one side is matched with an acceptable agent on the other side and no two agents of different sides would prefer to be matched to each other rather than to stick to their current situation. For marriage markets, this stability notion is known to be equivalent to core stability.

A matching is in the core if there is no subset of agents who, by forming only partnerships among themselves, can all obtain a strictly preferred outcome. Gale and Shapley (1962) show that the core of a marriage problem is non-empty. Although elements of the core have the property that they are stable once reached, it depends on the underlying environment whether it is possible to reach some core element from any initial situation. Stable sets as defined in von Neumann and Morgenstern (1944) address this concern.

A stable set is a set of outcomes that satisfies internal and external stability. As argued by von Neumann and Morgenstern (1944, p. 41), a stable set describes the "established order of society" or "accepted standard of behavior". Internal stability "expresses the fact that the standard of behavior is free from inner contradictions". External stability "can be used to discredit any non-conforming procedure".

vNM stable sets are crucially dependent on the concept of dominance. Under the standard definition, a matching is dominated by another matching if there is a coalition such that all its members prefer the latter matching to the former and no coalition member has a partner outside the coalition. A set of matchings is a vNM stable set if it satisfies the conditions of internal and external stability with respect to this dominance relation. Internal stability requires that no matching inside the set is dominated by a matching belonging to the set. External stability imposes that any matching outside the set is dominated by some matching belonging to the set. Ehlers (2007) shows that for one-to-one matching problems, the set of matchings in the core is a subset of any vNM stable set and a vNM stable set can contain matchings outside the core. Wako (2010) shows that the vNM stable set exists and is unique.

The standard dominance relation used to define vNM stable sets violates the assumption of coalitional sovereignty, the property that an objecting coalition cannot enforce the organization of

* Corresponding author.
E-mail addresses: P.Herings@maastrichtuniversity.nl (P. Jean-Jacques Herings),
ana.mauleon@usaintlouis.be (A. Mauleon), vincent.vannetelbosch@uclouvain.be
(V. Vannetelbosch).

agents outside the coalition. Such a violation is surprising since coalitional sovereignty is a very natural property to require when defining enforceability. For one-to-one matching problems, if a coalition deviates, then it is free to form any match between its members; it cannot affect existing matches between agents outside the coalition, and previous matches between coalition and non-coalition members are destroyed. The requirement that the coalition cannot affect existing matches between agents outside the coalition is violated in the standard definition of a vNM stable set.

Coalitional sovereignty is a natural requirement and has been imposed in other streams in the literature. In the literature on coalition formation, Hart and Kurz (1983) propose the γ and the δ model of coalition formation. Let some coalition structure be given and suppose some coalition deviates. Unaffected players are those players who are not part of the deviating coalition and were not together with any player of the deviating coalition in the original coalition structure. Coalitional sovereignty requires that nothing changes for the unaffected players and both the γ and the δ model respect coalitional sovereignty. Another issue is what happens to residual players, that is those players who were together with some player of the deviating coalition in the original coalition structure. The γ model assumes that the residual players become singletons after the deviation and the δ model assumes that they stay together. Kóczy and Lauwers (2004) study the accessibility of the core of a TU-game in coalitional form and also emphasize the importance of coalitional sovereignty. They call this property outsider independence.

For many-to-many matching problems, several authors have proposed and studied solution concepts that respect coalitional sovereignty, see in particular Echenique and Oviedo (2006) and Konishi and Ünver (2006). The important issue in many-to-many matching problems is not so much coalitional sovereignty, which is naturally assumed, but rather what happens to links between members of the deviating coalition and players outside that coalition, a problem closely related to the treatment of residual players in models of coalition formation.

An important stream in the literature on matching markets studies whether a decentralized process of successive blocking leads to a stable matching, see e.g. Roth and Vande Vate (1990), Klaus and Klijn (2007), and Kojima and Ünver (2008) for two-sided matching problems, and Chung (2000) and Diamantoudi et al. (2004) for roommate problems. All these papers formulate dominance relations that satisfy coalitional sovereignty.

Finally, several papers on the vNM stable set in the case of farsighted agents have used notions of enforceability that respect coalitional sovereignty, see Diamantoudi and Xue (2003) for hedonic games, Mauleon et al. (2011) for one-to-one matching problems, Klaus et al. (2011) for roommate markets, and Ray and Vohra (2015) for non-transferable utility games.

A further criticism of the standard definition of the vNM stable set is that it does not take into account that a deviation by a coalition can be followed by further deviations. This corresponds to the well-known critique by Harsanyi (1974) to the vNM stable set. Ray and Vohra (2015) emphasize that the notion of enforceability is especially delicate in the context of farsightedness as formulated by Harsanyi (1974). We argue here that, even in the standard myopic case, the same issue comes up when we apply the vNM stable set to one-to-one matching problems. We will follow the approach by van Deemen (1991) and Page and Wooders (2009), which takes into account that if a matching is blocked by some coalition and the resulting matching is not in the stable set itself, then further deviations will take place. This observation leads van Deemen (1991) to define the generalized stable set for abstract systems and Page and Wooders (2009) to define the stable set with respect to path dominance. We show by means of a simple example that not allowing for path dominance in the definition of the vNM stable set for one-to-one matching problems leads to highly undesirable conclusions when coalitional sovereignty is required.

Requiring coalitional sovereignty (C) and using the path dominance relation (P) to define internal and external stability, now referred to as CP internal stability and CP external stability, leads to the concept of the CP vNM set. Since in matching theory it is often assumed that only pairwise deviations are feasible, we also define the concept of the pairwise CP vNM set in an analogous way.

We show that there is a unique CP vNM set and a unique pairwise CP vNM set and that both sets coincide with the core. Although, as shown by Ehlers (2007), the core may not be a vNM stable set under the standard definition of the direct dominance relation,[1] it turns out to be the unique prediction when coalitional sovereignty and path dominance are taken into account.

Since the CP vNM set and the pairwise CP vNM set are based on paths of deviations resulting from the direct dominance relation, they should be thought of as myopic concepts. An alternative would be a farsighted approach based on the indirect dominance relation as introduced in Harsanyi (1974) and further developed in Chwe (1994). The vNM farsighted stable sets have been characterized in Mauleon et al. (2011) as the singleton core elements. Diamantoudi and Xue (2003) show that, in hedonic games with strict preferences, core partitions are farsightedly stable for a conservative notion of the concept.

The paper is organized as follows. Section 2 introduces one-to-one matching problems and standard notions of stability. Section 3 defines and characterizes both the CP vNM and the pairwise CP vNM set. Section 4 concludes and discusses directions for future research.

2. One-to-one matching problems

A one-to-one matching problem consists of a finite set of individuals N, partitioned into a set of men M and a set of women W. The set of non-empty subsets of N is denoted by $\mathcal{N}$. Each individual $i \in N$ has a complete and transitive *preference ordering* $\succ_i$ over the agents on the other side of the market and the prospect of being alone. Preferences are assumed to be strict. Let $\succ = ((\succ_m)_{m \in M}, (\succ_w)_{w \in W})$ be a preference profile. A one-to-one matching problem is a triple $(M, W, \succ)$.

A matching is a function $\mu : N \rightarrow N$ satisfying the following properties:

(i) For every $m \in M$, $\mu(m) \in W \cup \{m\}$.
(ii) For every $w \in W$, $\mu(w) \in M \cup \{w\}$.
(iii) For every $i \in N$, $\mu(\mu(i)) = i$.

The set of all matchings is denoted by $\mathcal{M}$. Given a matching $\mu \in \mathcal{M}$, individual $i \in N$ is said to be unmatched or single if $\mu(i) = i$. A matching μ is individually rational if each agent is acceptable to his or her mate, so for every $i \in N$ it holds that $\mu(i) \succ_i i$ or $\mu(i) = i$. A matching μ that is not individually rational can be blocked by any individual with an unacceptable partner. For a given matching μ, a pair $\{m, w\}$ is said to form a blocking pair if m and w are not matched to one another but prefer one another to their mates at μ, i.e. $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$. A matching μ is stable if it is not blocked by any individual or any pair of agents.

For every $i \in N$, we extend the preference ordering $\succ_i$ over the agent's potential partners to the set of matchings in the following

[1] There are not so many classes of games where the core is the unique vNM stable set of the game. One example of such a class is the class of convex games, see Shapley (1971).

way. We say that agent i prefers the matching μ' to the matching μ if $\mu'(i) \succ_i \mu(i)$ and we write $\mu' \succ_i \mu$. A coalition $S \in \mathcal{N}$ is said to block a matching $\mu \in \mathcal{M}$ if there exists a matching $\mu' \in \mathcal{M}$ such that $\mu'(S) = S$ and $\mu' \succ_S \mu$, where $\mu' \succ_S \mu$ is defined as $\mu'(i) \succ_i \mu(i)$ for every $i \in S$. Such a matching μ' is said to dominate the matching μ via S. The set of all such matchings is denoted by $f_S(\mu)$, so

$$f_S(\mu) = \{\mu' \in \mathcal{M} \mid \mu'(S) = S \text{ and } \mu' \succ_S \mu\}.$$

We say that the matching μ' dominates the matching μ if there is a coalition $S \in \mathcal{N}$ such that μ' dominates the matching μ via S and we denote the set of all such matchings by $f(\mu)$, so

$$f(\mu) = \bigcup_{S \in \mathcal{N}} f_S(\mu).$$

The core C of the matching problem $(M, W, \succ)$ consists of all matchings that are not blocked by any coalition, so

$$C = \{\mu \in \mathcal{M} \mid f(\mu) = \emptyset\}.$$

It has been shown by Gale and Shapley (1962) that the core of a matching problem is non-empty. Also, a matching is stable if and only if it is not blocked by a coalition of size one or two and if and only if it belongs to the core, see Theorem 3.3 in Roth and Sotomayor (1990).

We next present the standard definition of the von Neumann Morgenstern (vNM) stable set as introduced in von Neumann and Morgenstern (1944) when applied to matching problems.

Definition 1. A set of matchings $V \subset \mathcal{M}$ is a vNM stable set if it satisfies the following:

(i) Internal stability: For every $\mu, \mu' \in V$, it holds that $\mu' \notin f(\mu)$.
(ii) External stability: For every $\mu \in \mathcal{M} \setminus V$, it holds that $f(\mu) \cap V \neq \emptyset$.

A set of matchings is a vNM stable set if it satisfies internal and external stability. Internal stability requires that no matching inside the vNM stable set is dominated by a matching belonging to the vNM stable set. External stability states that any matching outside the vNM stable set is dominated by some matching that belongs to the vNM stable set.

As for the definition of the core, there is no loss of generality to restrict attention to coalitions of size one and two when formulating internal and external stability or to restrict attention to pairwise deviations. It is immediate that the core is contained in any vNM stable set. However, since the core may not satisfy external stability, it may not be a vNM stable set itself.

Given two matchings $\mu, \mu' \in \mathcal{M}$, let $\mu \vee \mu' : N \to N$ be the function defined by

$$(\mu \vee \mu')(m) = \max_{\succ_m}\{\mu(m), \mu'(m)\}, \quad m \in M,$$
$$(\mu \vee \mu')(w) = \min_{\succ_w}\{\mu(w), \mu'(w)\}, \quad w \in W.$$

Similarly, we define $\mu \wedge \mu' : N \to N$ as

$$(\mu \wedge \mu')(m) = \min_{\succ_m}\{\mu(m), \mu'(m)\}, \quad m \in M,$$
$$(\mu \wedge \mu')(w) = \max_{\succ_w}\{\mu(w), \mu'(w)\}, \quad w \in W.$$

In general, the functions $\mu \vee \mu'$ and $\mu \wedge \mu'$ need not be matchings.

A set $X \subset \mathcal{M}$ of matchings is a lattice if for every $\mu, \mu' \in X$ it holds that $\mu \vee \mu' \in X$ and $\mu \wedge \mu' \in X$. The set X is a distributive lattice if X is a lattice and for every $\mu, \mu', \mu'' \in X$ it holds that

$$\mu \vee (\mu' \wedge \mu'') = (\mu \vee \mu') \wedge (\mu \vee \mu''),$$
$$\mu \wedge (\mu' \vee \mu'') = (\mu \wedge \mu') \vee (\mu \wedge \mu'').$$

Knuth (1976) shows that the core C of a matching problem is a distributive lattice. Ehlers (2007) shows that V is a vNM stable

set only if V is a maximal set satisfying $C \subset V$, V is a distributive lattice, and the set of single agents is the same for every element of V. Moreover, V is a vNM stable set if V is the unique maximal set satisfying $C \subset V$, V is a distributive lattice, and the set of single agents is the same for every element of V. A vNM stable set V contains the core but can also contain matchings outside the core. Wako (2010) shows the existence and the uniqueness of the vNM stable set.

3. vNM sets respecting coalitional sovereignty

The standard dominance relation used in the definition of the vNM stable sets in one-to-one matching problems denies the coalitional sovereignty of players outside the objecting coalition and this could have undesirable consequences. Since in the formulation of the vNM stable set using the dominance relation f, a coalition $S \in \mathcal{N}$ can impose any matching as long as its restriction to S is feasible for S, the objecting coalition is given implicitly the power to arrange the matchings of players outside the coalition, which is clearly undesirable. Ray and Vohra (2015) express the same criticism towards the vNM stable set for non-transferable utility games.

In the following, we adapt the dominance relation in such a way that coalitional sovereignty of players outside the objecting coalition is respected. Contrary to the dominance relation f, it may now matter whether the size of the objecting coalition is unrestricted or takes place by individuals and blocking pairs. We therefore formulate two dominance relations.

Assume, first, that we do not restrict the size of the objecting coalition. In order to respect coalitional sovereignty, the blocking of a matching $\mu \in \mathcal{M}$ by a coalition $S \in \mathcal{N}$ should result in a matching μ' such that the players $i \in \mu(S) \setminus S$ become single, whereas the players in $N \setminus (S \cup \mu(S))$ are not affected by this blocking. We therefore restrict attention to matchings μ' with this feature and define the set of matchings that dominate μ via S as follows:

$$g_S(\mu) = \{\mu' \in \mathcal{M} \mid \mu'(S) = S,$$
$$\text{for every } i \in \mu(S) \setminus S, \; \mu'(i) = i,$$
$$\text{for every } i \in N \setminus (S \cup \mu(S)),$$
$$\mu'(i) = \mu(i), \mu' \succ_S \mu\}.$$

It clearly holds that $g_S(\mu) \subset f_S(\mu)$. The set of matchings that dominate μ is now defined as follows:

$$g(\mu) = \bigcup_{S \in \mathcal{N}} g_S(\mu).$$

It clearly holds that $g(\mu) \subset f(\mu)$. We refer to an element of $g(\mu)$ as an improvement.

In matching theory, it is typical to restrict blocking to individuals and blocking pairs. Let some matching $\mu \in \mathcal{M}$ be given. An individual $i \in N$ can destroy his match with player $\mu(i)$ whenever $\mu(i) \neq i$. The players in $N \setminus \{i, \mu(i)\}$ are not affected by the destruction of this match, so remain matched to the same partner. The resulting matching is denoted by $\mu' = \mu - (i, \mu(i))$.

A blocking pair (m, w) can destroy their existing matches, if any, and form a match between them. The players in $N \setminus \{m, w, \mu(m), \mu(w)\}$ are not affected by the formation of such a match, so keep their existing partners. The resulting matching is denoted by $\mu' = \mu + (m, w)$.

Let some matching $\mu \in \mathcal{M}$ and some coalition $S \in \mathcal{N}$ with $S = \{i\}$ for some $i \in N$ or $S = \{m, w\}$ with $m \in M$ and $w \in W$ be given. We define the set of matchings that dominate μ via S as follows:

$$h_S(\mu) = \begin{cases} \{\mu' \in \mathcal{M} \mid \mu' = \mu - (i, \mu(i)) \text{ and } i \succ_i \mu(i)\}, & S = \{i\}, \\ \{\mu' \in \mathcal{M} \mid \mu' = \mu + (m, w), \\ \quad w \succ_m \mu(m), \text{ and } m \succ_w \mu(w)\}, & S = \{m, w\}. \end{cases}$$

Observe that for singleton coalitions S it holds that $h_S(\mu) = g_S(\mu)$ and for coalitions S consisting of a man and a woman we have $h_S(\mu) \subset g_S(\mu)$. In the latter case, inequality may result because under h_S it is not allowed that m and w destroy their current match without forming a new one together. Notice that in all cases, $h_S(\mu)$ is either the empty set or a singleton. We define the set of matchings that dominate μ as follows:

$$h(\mu) = \left(\bigcup_{i \in N} h_{\{i\}}(\mu) \right) \cup \left(\bigcup_{(m,w) \in M \times W} h_{\{m,w\}}(\mu) \right).$$

It clearly holds that $h(\mu) \subset g(\mu)$. We refer to an element of $h(\mu)$ as a pairwise improvement.

Although it might be tempting to define the vNM stable set by replacing $f(\mu)$ by either $g(\mu)$ or $h(\mu)$ in Definition 1, the following example shows this to be problematic.

Example 1. Consider a matching problem $(M, W, \succ)$, where $M = \{m_1, m_2\}$, $W = \{w_1, w_2\}$, and the preferences of the individuals are as follows.

m_1	m_2	w_1	w_2
w_1	w_1	m_1	m_2
w_2	w_2	m_2	m_1
m_1	m_2	w_1	w_2.

We represent a matching μ by the set of (m, w) pairs that are matched in μ. There is a single stable matching $\mu^* = \{(m_1, w_1), (m_2, w_2)\}$. The core has μ^* as its unique element. For every $\mu \in \mathcal{M}$ such that $\mu \neq \mu^*$ we have that $\mu^* \in f(\mu)$. It now follows easily that there is a unique vNM stable set V_f and that this set is equal to the core, $V_f = C = \{\mu^*\}$.

However, according to the dominance relation g, the matching μ^* does not dominate the matching $\mu' = \{(m_1, w_2), (m_2, w_1)\}$. Indeed, only the grand coalition N can enforce the matching μ^* from μ' under g. Since $\mu' \succ_{m_2} \mu^*$, m_2 would block the deviation by the grand coalition from μ' to μ^*.

Let V_g be a vNM stable set based on the dominance relation g. Since $g(\mu^*) = \emptyset$, external stability requires $\mu^* \in V_g$. For every matching μ not equal to μ^*, μ', or $\{(m_2, w_1)\}$, it holds that $\mu^* \in g(\mu)$, so in order not to violate internal stability, $\mu \notin V_g$. Since $g(\mu')$ only contains the matching $\{(m_1, w_1)\}$ and $\{(m_1, w_1)\} \notin V_g$, it holds that $\mu' \in V_g$ by external stability. Since $\mu' \in g(\{(m_2, w_1)\})$, it follows that $V_g = \{\mu^*, \mu'\}$.

The fact that μ' belongs to V_g is problematic. The only improvement from μ' is given by the matching $\{(m_1, w_1)\}$. Since $\{(m_1, w_1)\}$ does not belong to V_g, one needs μ' to be part of V_g in order to satisfy external stability. However, starting from the matching $\{(m_1, w_1)\}$, there is only one further improvement, i.e. to μ^*, and from μ^* there are no further improvements. Given a matching μ', one would therefore expect to end up in μ^*, which is part of V_g, and μ' should not be regarded as stable.

When using the dominance relation h in Definition 1, the sets $\{\mu^*\}$ and $\{\mu^*, \mu'\}$ would not qualify as stable sets. Denoting the matching where every individual is unmatched by $\mu^\emptyset$, we have that $\mu^* \notin h(\mu^\emptyset)$ and $\mu' \notin h(\mu^\emptyset)$, so external stability would be violated.

Let V_h be a vNM stable set based on the dominance relation h. Since $h(\mu^*) = \emptyset$, external stability requires $\mu^* \in V_h$. Since $\mu^* \in h(\{(m_1, w_1)\})$ and $\mu^* \in h(\{(m_2, w_2)\})$, it follows from internal stability that $\{(m_1, w_1)\} \notin V_h$ and $\{(m_2, w_2)\} \notin V_h$.

Since $h(\mu')$ contains only the matching $\{(m_1, w_1)\}$ and $\{(m_1, w_1)\} \notin V_h$, it follows from external stability that $\mu' \in V_h$.

The set $h(\mu^\emptyset)$ consists of the two matchings $\{(m_1, w_1)\}$ and $\{(m_2, w_2)\}$. Since $\{(m_1, w_1)\} \notin V_h$ and $\{(m_2, w_2)\} \notin V_h$, external stability then requires that $\mu^\emptyset \in V_h$.

Finally, since $\mu' \in h(\{(m_1, w_2)\})$ and $\mu' \in h(\{(m_2, w_1)\})$, it follows from internal stability that $\{(m_1, w_2)\} \notin V_h$ and $\{(m_2, w_1)\} \notin V_h$.

We reach the conclusion that $V_h = \{\mu^*, \mu', \mu^\emptyset\}$ is the unique vNM stable set based on the dominance relation h. Since $\mu^\emptyset$ is the worst matching for all individuals, this is a highly undesirable conclusion. $\square$

Example 1 shows that the assumption of coalitional sovereignty affects the vNM stable set. Also, under coalitional sovereignty, it matters whether blocking coalitions can be of any size or if blocking is restricted to pairwise deviations and deviations by individuals. The reason for this difference is that external stability requires each matching not in the stable set to be directly dominated by an element in the stable set. If a matching μ is dominated by another matching μ' via a coalition S of cardinality greater than two, then it may not be the case that μ' can be enforced from μ by an individual or a pair.

Rather than requiring direct domination, it should be sufficient for external stability that there is a sequence of matchings, each matching in the sequence dominating the previous one, that ultimately reaches a matching in the stable set. Indeed, if a matching that is not in the stable set is blocked by some coalition and the resulting matching is not stable itself, then further improvements will take place. We make a similar requirement for testing internal stability. That is, if a matching belongs to the stable set, then there should be no sequence of matchings, each matching in the sequence being dominated by the next one, that ultimately reaches another matching in the stable set. These requirements correspond to path dominance as defined in van Deemen (1991) and Page and Wooders (2009).

Let some $\mu \in \mathcal{M}$ be given. The set

$$h^2(\mu) = h(h(\mu)) = \{\mu'' \in \mathcal{M} \mid \exists \mu' \in h(\mu) \text{ such that } \mu'' \in h(\mu')\}$$

consists of those matchings that can be reached by a composition of two pairwise improvements starting at μ. We extend this definition and, for $k \in \mathbb{N}$, we define $h^k(\mu)$ as those matchings that can be reached from μ by a composition of k pairwise improvements. The transitive closure of h is denoted by h^∞ and is defined as

$$h^\infty(\mu) = \bigcup_{k \in \mathbb{N}} h^k(\mu), \quad \mu \in \mathcal{M}.$$

Since the set $\mathcal{M}$ is finite, it holds that, for some $k' \in \mathbb{N}$, for every $\mu \in \mathcal{M}$, $h^\infty(\mu) = \bigcup_{k \in \{1, \dots, k'\}} h^k(\mu)$.

When taking into account the requirements of coalitional sovereignty (C) and path dominance (P), while only allowing for pairwise deviations, we obtain the notion of a pairwise CP vNM set, more formally defined as follows.

Definition 2. A set of matchings $V \subset \mathcal{M}$ is a pairwise CP vNM set if it satisfies the following:

(i) Pairwise CP internal stability: For every $\mu, \mu' \in V$, it holds that $\mu' \notin h^\infty(\mu)$.
(ii) Pairwise CP external stability: For every $\mu \in \mathcal{M} \setminus V$, it holds that $h^\infty(\mu) \cap V \neq \emptyset$.

Condition (i) of Definition 2 corresponds to pairwise CP internal stability. For any two matchings μ and μ' in the pairwise CP vNM set V it does not hold that there is a sequence of improvements from one to the other. Condition (ii) of Definition 2 expresses pairwise CP external stability. For every matching μ outside the pairwise CP vNM set V it holds that there is $\mu' \in V$ and a sequence of matchings $\mu_0, \dots, \mu_{k'}$ with $\mu_0 = \mu$ and $\mu_{k'} = \mu'$ such that, for every $k \in \{0, \dots, k' - 1\}$, $\mu_{k+1} \in h(\mu_k)$.

The definition of a pairwise CP vNM set is equivalent to the definition of a generalized stable set for abstract systems as

introduced in van Deemen (1991), when we consider the abstract system $(\mathcal{M}, h)$. It corresponds to the stable set with respect to path dominance in Page and Wooders (2009) when h is taken to be the direct dominance relation.

Theorem 1. *A set of matchings is a pairwise CP vNM set if and only if it is equal to the core.*

Proof. We start with the "if" part of the proof and show that the core C is a pairwise CP vNM set.

Let some $\mu \in C$ be given. It holds by definition of the core that $f(\mu) = \emptyset$. Since $h(\mu) \subset f(\mu)$, it follows that $h(\mu) = \emptyset$. Pairwise CP internal stability follows.

Let some $\mu \in \mathcal{M} \setminus C$ be given. We have to show that $h^\infty(\mu) \cap C \neq \emptyset$. It follows from Theorem in Roth and Vande Vate (1990) that $h^\infty(\mu)$ contains a stable matching. Since a stable matching belongs to C by Theorem 3.3 in Roth and Sotomayor (1990), pairwise CP external stability follows.

We continue with the "only if" part of the proof and show that C is the only pairwise CP vNM set. Let V be a pairwise CP vNM set. For every $\mu \in C$ it holds that $h(\mu) \subset f(\mu) = \emptyset$. Since V satisfies pairwise CP external stability, it follows that $C \subset V$.

Suppose $V \setminus C \neq \emptyset$. Let some $\mu' \in V \setminus C$ be given. Combining again theorem in Roth and Vande Vate (1990) and Theorem 3.3 in Roth and Sotomayor (1990), it follows that $h^\infty(\mu') \cap C \neq \emptyset$, therefore $h^\infty(\mu') \cap V \neq \emptyset$, thereby violating pairwise CP internal stability. Consequently, it holds that $V \setminus C = \emptyset$. ∎

Theorem of Roth and Vande Vate (1990) proposes a family of random processes, beginning from an arbitrary matching and selecting a blocking pair at random to create a new matching, that will reach a stable matching with probability one. In the proof of Theorem 1, we use this result to prove that the core satisfies pairwise CP external stability and that the core is the unique pairwise CP vNM set.

An alternative to the pairwise CP vNM set is to allow a coalition of any size to form and block, while respecting coalitional sovereignty. We denote the transitive closure of g by g^∞.

Definition 3. A set of matchings $V \subset \mathcal{M}$ is a CP vNM set if it satisfies the following:

(i) CP internal stability: For every $\mu, \mu' \in V$, it holds that $\mu' \notin g^\infty(\mu)$.
(ii) CP external stability: For every $\mu \in \mathcal{M} \setminus V$, it holds that $g^\infty(\mu) \cap V \neq \emptyset$.

The definition of a CP vNM set is equivalent to the definition of a generalized stable set for abstract systems as introduced in van Deemen (1991), when we consider the abstract system $(\mathcal{M}, g)$ and it corresponds to the stable set with respect to path dominance in Page and Wooders (2009) when g is taken to be the direct dominance relation.

Let $\mu \in \mathcal{M}$ be a matching. We have already argued that $h(\mu) \subset g(\mu)$. If it would be true that $h^\infty(\mu) = g^\infty(\mu)$ then it would be immediate that the result of Theorem 1 is also valid for the CP vNM set. The following example demonstrates that $h^\infty(\mu)$ can be a proper subset of $g^\infty(\mu)$.[2]

Example 2. Consider a matching problem $(M, W, \succ)$, where $M = \{m_1, m_2\}$, $W = \{w_1, w_2\}$, and the preferences of the individuals are as follows.

m_1	m_2	w_1	w_2
w_1	w_2	w_1	w_2
w_2	w_1	m_1	m_2
m_1	m_2	m_2	m_1.

---

[2] We are grateful to an anonymous referee for suggesting this example.

The preferences in Example 2 are such that man m_1 prefers to be married to w_1 rather than to w_2 and m_2 prefers to be married to w_2 rather than to w_1. Also, woman w_1 prefers m_1 over m_2 and woman w_2 prefers m_2 over m_1, but contrary to the men, women prefer not to be married at all. As in Example 1, we define $\mu' = \{(m_1, w_2), (m_2, w_1)\}$ and $\mu^* = \{(m_1, w_1), (m_2, w_2)\}$. It is easily verified that $\mu^* \in g(\mu')$ and therefore $\mu^* \in g^\infty(\mu')$. We argue next that $\mu^* \notin h^\infty(\mu')$. The first observation is that at any $\mu \in h(\mu')$ it holds that one of the women is single. If μ is obtained from μ' by the destruction of a match by an individual, this is obviously true. If μ is obtained from μ' by a blocking pair (m, w), then $\mu'(m)$ is single at μ. Observe next that once a woman w is single, she has reached her most preferred situation, and can no longer be involved in a further pairwise improvement. It follows that $\mu^* \notin h^\infty(\mu')$ and therefore $h^\infty(\mu')$ is a proper subset of $g^\infty(\mu')$.

The next result demonstrates that $g^\infty(\mu)$ and $h^\infty(\mu)$ are identical for individually rational matchings μ.

Theorem 2. *For every individually rational $\mu \in \mathcal{M}$, it holds that $g^\infty(\mu) = h^\infty(\mu)$.*

Proof. Let some $\mu \in \mathcal{M}$ be given. Obviously, it holds that $h^\infty(\mu) \subset g^\infty(\mu)$.

To show the converse, it is sufficient to show that for every $\mu \in \mathcal{M}$, for every $\mu' \in g(\mu)$, there is a sequence $\mu_0, \ldots, \mu_{k'}$ such that $\mu = \mu_0$, $\mu' = \mu_{k'}$ and, for every $k = 0, \ldots, k' - 1$, $\mu_{k+1} \in h(\mu_k)$, so $\mu' \in h^{k'}(\mu)$.

Let $\mu \in \mathcal{M}$, $\mu' \in g(\mu)$, and $S \in \mathcal{N}$ be given such that $\mu'(S) = S$, for every $i \in \mu(S) \setminus S$, $\mu'(i) = i$, for every $i \in N \setminus (S \cup \mu(S))$, $\mu'(i) = \mu(i)$, and $\mu' \succ_S \mu$. Since $\mu' \succ_S \mu$ and μ is individually rational, it holds that, for every $i \in S$, $\mu'(i) \succ_i \mu(i) \succeq_i i$, so in particular $\mu'(i) \neq i$. Therefore, there is k' such that $S = \{i_0, \ldots, i_{2k'-1}\}$ and for every $i_k \in S$ with $k \leq k' - 1$, $\mu'(i_k) = i_{k+k'}$.

We form the sequence $\mu_0, \ldots, \mu_{k'}$ by defining, for $k = 0, \ldots, k' - 1$, $\mu_{k+1} = \mu_k + (i_k, i_{k+k'})$. Let some $k = 0, \ldots, k' - 1$ be given. It holds that

$$i_{k+k'} = \mu'(i_k) \succ_{i_k} \mu_k(i_k),$$

since $\mu_k(i_k) = i_k$ or $\mu_k(i_k) = \mu(i_k)$. Similarly, it follows that $i_k \succ_{i_{k+k'}} \mu_k(i_{k+k'})$. We have shown that $\mu_{k+1} \in h(\mu_k)$. ∎

We now show the analogue of Theorem 1 for CP vNM sets.

Theorem 3. *A set of matchings is a CP vNM set if and only if it is equal to the core.*

Proof. We start with the "if" part of the proof and show that the core C is a CP vNM set.

Let some $\mu \in C$ be given. It holds by definition of the core that $f(\mu) = \emptyset$. Since $g(\mu) \subset f(\mu)$, it follows that $g(\mu) = \emptyset$. CP internal stability follows.

Let some $\mu \in \mathcal{M} \setminus C$ be given. We have to show that $g^\infty(\mu) \cap C \neq \emptyset$. Since $h^\infty(\mu) \subset g^\infty(\mu)$, this follows from pairwise CP external stability as shown in Theorem 1.

We continue with the "only if" part of the proof and show that C is the only CP vNM set. Let V be a CP vNM set. For every $\mu \in C$ it holds that $g(\mu) \subset f(\mu) = \emptyset$. Since V satisfies CP external stability, it follows that $C \subset V$.

Suppose $V \setminus C \neq \emptyset$. Let some $\mu' \in V \setminus C$ be given. Since $g^\infty(\mu') \cap C \neq \emptyset$ by CP external stability of C, it holds that $g^\infty(\mu') \cap V \neq \emptyset$, thereby violating CP internal stability. Consequently, it holds that $V \setminus C = \emptyset$. ∎

In general, the vNM stable set with g as the dominance relation, V_g, is substantially different from the vNM stable set with h as the dominance relation, V_h, as shown in Example 1. However, these differences disappear whenever the external stability as based on direct dominance is replaced by path dominance.

4. Conclusion

We study stable sets for one-to-one matching problems that take the requirement of coalitional sovereignty into account. Coalitional sovereignty means that the objecting coalition does not have the power to arrange the matchings of agents outside the coalition. In these new notions of stable sets, the internal and external stability conditions are formulated using path dominance, following van Deemen (1991) and Page and Wooders (2009), and motivated by Example 1.

We consider both general coalitional deviations as well as pairwise deviations, resulting in the concepts of CP vNM set and pairwise CP vNM set. We obtain a unique prediction for both concepts: the matchings that belong to the core. This contrasts with the findings of Ehlers (2007), who shows that the set of matchings in the core is a subset of any vNM stable set and may contain matchings outside the core when the dominance relation does not respect coalitional sovereignty.

In the characterization of the pairwise CP vNM set and the CP vNM set, Theorems 1 and 3, we use the result of Roth and Vande Vate (1990) that starting from an arbitrary matching it is possible to find a sequence of blocking pairs that ultimately reaches a stable matching. We use this result to prove that the core satisfies pairwise CP external stability and that the core is the unique (pairwise) CP vNM set.

We have restricted our analysis to one-to-one matching problems, but the concepts of CP vNM set and pairwise CP vNM set can also be defined for roommate problems or many-to-many matching problems.[3] For roommate problems, Chung (2000) and Diamantoudi et al. (2004) provide conditions on preferences such that a decentralized process along the lines of Roth and Vande Vate (1990) converges to a stable roommate matching with probability one. Under such conditions on preferences, one can show the CP vNM set and the pairwise CP vNM set to be unique and to coincide with the set of stable roommate matchings, following similar arguments as in the proofs of Theorems 1 and 3.

Since for roommate problems, the notions of CP vNM set and pairwise CP vNM set correspond to generalized stable sets for suitably chosen abstract systems in the sense of van Deemen (1991), we obtain existence of CP vNM sets and pairwise CP vNM sets without conditions on preferences. Moreover, such sets consists of one representative element of each closed cycle in the sense of Jackson and Watts (2002). An interesting direction for future research would be whether it is possible to characterize these closed cycles.

In a similar way, using the results of Klaus and Klijn (2007), our results can be extended to matching problems with couples and, using the results of Kojima and Ünver (2008), to many-to-many matching problems. In these more challenging environments, there is again a need for assumptions on the preferences to obtain a neat characterization of the (pairwise) CP vNM set. The concepts exist without conditions on the preferences. For the many-to-many case, there are several possibilities to define dominance correspondences that satisfy coalitional sovereignty, depending on the treatment of matches between members of the deviating coalition and non-coalition members. Finally, contrary to the one-to-one matching case, for the many-to-many case, it matters whether one considers only pairwise deviations or not.

References

Chung, K.-S., 2000. On the existence of stable Roommate matchings. Games Econom. Behav. 33, 206–230.

Chwe, M.S.-Y., 1994. Farsighted coalitional stability. J. Econom. Theory 63, 299–325.

Diamantoudi, E., Miyagawa, E., Xue, L., 2004. Random paths to stability in the Roommate problem. Games Econom. Behav. 48, 18–28.

Diamantoudi, E., Xue, L., 2003. Farsighted stability in Hedonic games. Soc. Choice Welf. 21, 39–61.

Echenique, F., Oviedo, J., 2006. A theory of stability in many-to-many matching markets. Theor. Econ. 1, 233–273.

Ehlers, L., 2007. Von Neumann-Morgenstern stable sets in matching problems. J. Econom. Theory 134, 537–547.

Gale, D., Shapley, L.S., 1962. College admissions and the stability of marriage. Amer. Math. Monthly 69, 9–15.

Harsanyi, J.C., 1974. An equilibrium-point interpretation of stable sets and a proposed alternative definition. Manage. Sci. 20, 1472–1495.

Hart, S., Kurz, M., 1983. Endogenous formation of coalitions. Econometrica 51, 1047–1064.

Jackson, M.O., Watts, A., 2002. The evolution of social and economic networks. J. Econom. Theory 106, 265–295.

Klaus, B., Klijn, F., 2007. Paths to stability for matching markets with couples. Games Econom. Behav. 58, 154–171.

Klaus, B., Klijn, F., Walzl, M., 2011. Farsighted stability for Roommate markets. J. Public Econ. Theory 13, 921–933.

Knuth, D.E., 1976. Marriages Stables. Les Presses de l'Université de Montreal, Montreal.

Kóczy, L.Á., Lauwers, L., 2004. The coalition structure core is accessible. Games Econom. Behav. 48, 86–93.

Kojima, F., Ünver, M.U., 2008. Random paths to pairwise stability in many-to-many matching problems: A study on market equilibration. Internat. J. Game Theory 36, 473–488.

Konishi, H., Ünver, M.U., 2006. Credible group stability in many-to-many matching problems. J. Econom. Theory 129, 57–80.

Mauleon, A., Vannetelbosch, V., Vergote, W., 2011. Von Neumann-Morgenstern farsightedly stable sets in two-sided matching. Theor. Econ. 6, 499–521.

Page Jr., F.H., Wooders, M., 2009. Strategic basins of attraction, the path dominance core, and network formation games. Games Econom. Behav. 66, 462–487.

Ray, D., Vohra, R., 2015. The farsighted stable set. Econometrica 83, 977–1011.

Roth, A.E., Sotomayor, M.A.O., 1990. Two-sided Matching, A Study in Game-theoretic Modeling and Analysis. In: Econometric Society Monographs, vol. 18. Cambridge University Press, Cambridge.

Roth, A.E., Vande Vate, J.H., 1990. Random paths to stability in two-sided matching. Econometrica 58, 1475–1480.

Shapley, L.S., 1971. Cores of convex games. Internat. J. Game Theory 1, 11–26.

van Deemen, A.M.A., 1991. A note on generalized stable sets. Soc. Choice Welf. 8, 255–260.

von Neumann, J., Morgenstern, O., 1944. Theory of Games and Economic Behavior. Princeton University Press, Princeton, New Jersey.

Wako, J., 2010. A polynomial-time algorithm to find von Neumann-Morgenstern stable matchings in marriage games. Algorithmica 58, 188–220.

---

[3] We thank an anonymous referee for making us emphasize the greater generality of our results.