Farsighted stability with heterogeneous expectations*

Francis Bloch† Anne van den Nouweland‡

March 22, 2017

Abstract

This paper analyzes farsighted stable sets when agents have heterogeneous expectations over the dominance paths. We consider expectations functions satisfying two properties of path-persistence and consistency. We show that farsighted stable sets always exist. Any singleton farsighted stable set with common expectations is a farsighted stable set with heterogeneous expectations. We characterize singleton farsighted stable sets with heterogeneous expectations in one-to-one matching and voting models, and show that the relaxation of the hypothesis of common expectations greatly expands the set of states which can be supported as singleton farsighted stable sets.

JEL Classification Numbers: C71, D72, D74

Keywords: farsighted stable sets, heterogeneous expectations, one-to-one matching, voting, effectivity functions

*We thank Bhaskar Dutta and Rajiv Vohra for stimulating discussions on this project. This work was supported by the Agence Nationale de la Recherche under contract ANR-13-BSHS1-0010.

†Université Paris 1 and Paris School of Economics, 48 Boulevard Jourdan, 75014 Paris, France. Email francis.bloch@univ-paris1.fr

‡Department of Economics 435 PLC, 1285 University of Oregon, Eugene, OR 97403-1285, USA. Email annev@oregon.edu

1 Introduction

Dominance is a key concept to assess the stability of cooperative relations and agreements. A situation a is said to dominate a situation b if there exists a coalition of players who can engineer the move from b to a and all prefer situation a to situation b. Ever since the seminal book by von Neumann and Morgenstern [13], dominance has been used to define cooperative solution concepts such as the core, the stable set or the bargaining set.[1] As Harsanyi [6] first pointed out, dominance implies a myopic behavior where agents only look one step ahead to assess the result of their actions. But if the situation they expect can itself be dominated by another situation, what should agents expect? Harsanyi [6] proposes to consider indirect dominance by sequences of situations to capture the farsighted expectations of agents: a situation a indirectly dominates a situation b if there exists a sequence of situations $a_0 = b, .., a_K = a$ such that, for any move along the sequence from a_k to a_{k+1} there exists a coalition of players who can engineer the move from a_k to a_{k+1} and who all prefer the final situation a to situation a_k. (Note that this formal definition of indirect dominance does not appear in Harsanyi [6] but in Chwe [2].)

The Harsanyi-Chwe definition of indirect dominance takes care of one objection to the original notion of dominance: agents now envision the whole sequence of moves yielding to stable situations. But the definition still suffers from several flaws. First, as was already acknowledged by Chwe [2], it assumes optimistic behavior on the part of agents, who may choose to leave a situation b because there exists *one* chain of dominance leading them to a situation they prefer. But there is no guarantee that this chain of dominance will indeed be chosen.[2] Second, as noted by Konishi and Ray [9] and Dutta and Vohra [4], this definition allows for inconsistent expectations. Different chains of dominance may imply different behaviors by the same player at some state.[3] Third, in some applications, the definition of situations and enforcement may allow for a coalition to choose arbitrarily the payoff of agents outside the coalition, creating chains of dominance where some agents receive very low payoffs chosen by other agents. As Ray and Vohra [14] notice, this may result in absurd predictions for TU games where situations are described by the distribution of a fixed resource over agents.[4]

In this paper, we tackle a different issue on the Harsanyi-Chwe definition of indirect dominance. We argue that the definition of an indirect dominance path imposes perfect

[1]The definition of dominance appears in von Neumann and Morgenstern [13], Section 30.1 p. 264.

[2]Chwe [2], Xue [17] and Mauleon and Vannetelbosch [10] propose solution concepts which deal with the problem: the largest consistent set for Chwe [2] , stability under the conservative standard of behavior for Xue [17] and the cautious largest consistent set for Mauleon and Vannetelbosch [10].

[3]This inconsistency disappears if one takes as a primitive description of agents' expectations an expectation function, assigning a final outcome to any state as in Jordan [8] or a path to any agent at any state as in Dutta and Vohra [4].

[4]Ray and Vohra [14] propose to remedy the problem by imposing a new condition of coalitional sovereignty, which prevents agents in a coalition to impose payments on agents outside the coalition.

coordination of the agents on the sequence of moves that are going to be played. In particular, if agents' expectations are described by expectation functions, as in Jordan [8] and Dutta and Vohra [4], all agents must share *common* expectations on the path of play. Our objective in this paper is to relax the assumption of perfect coordination of expectations, and to define and study farsighted behavior when agents entertain heterogeneous expectations on the dominance paths.

Our starting point is to define, for every individual, an expectation function associating to each situation the path of play that the agent expects from this point on. Clearly, we cannot allow for any expectation function and need to put restrictions on what individual agents may expect. Our first condition, called *path-persistence*, asserts that if an agent expects a path starting from a situation b to a situation a, they must expect the same continuation path to be followed after any situation in the path from b to a. In other words, expectations are persistent along a path, and agents cannot change their expectations somewhere along the path between b and a. This condition is reminiscent of conditions of inter-temporal consistency in the literature on dynamic optimization. It first guarantees that expectations are path-independent: the expectation from point b on are the same irrespective of the path that was follows up to b. This condition precludes agents from having different fixed limited windows of expectations: if agents believe that a path ends at a they cannot believe that there is another path starting from a.[5] The assumption of path-persistence is thus needed to define unambiguously the expectations of agents at any situation.

We also require that players' expectations be based on a rational behavior of the players. To this end, we use expectation functions to determine whether agents in a coalition prefer to move out of any given situation. If the answer is positive, we call a transition rationalizable. A profile of expectation functions is *consistent* if for any player at any situation the expected path is a succession of rationalizable moves. In other words, every agent constructs a theory of the moves which are going to be played, and this theory can be justified given the expectations of the other agents. Obviously, when agents have common expectations, consistency of the common expectation function collapses to a verification that all players are willing to move along the expected path.

When agents have different expectations, we cannot require that all agents agree on an indirect dominance path as this would surely imply that they share common expectations. This leaves us with an array of possible definitions of indirect dominance and stable sets. We choose to define dominance relative to the expectations of one player and characterize stable sets which can be supported by indirect dominance relative to the expectations of some player. In other words, a set of outcomes will be deemed stable if we can find one agent whose expectations support this set of outcomes as a stable set.

A profile of expectations where all agents believe that every situation is absorbent is

[5] By contrast, Herings, Mauleon and Vannetelbosch [7] study a situation where agents have limited foresight characterized by a fixed number of steps. They run into difficulties in describing agents' behavior when the length of dominance chains exceeds the number of steps of farsightedness.

clearly consistent and path-persistent. Given this profile of expectations, no move is rationalizable and hence the entire set of situations is always stable. This argument shows that existence of stable sets is not an issue in our model. Instead, we choose to focus attention on minimal stable sets, and our analysis concentrates on the existence of singleton stable sets, namely situations that can be attained from any other situation through an indirect dominance path. As our definition of indirect dominance does not require common expectations, it is more permissive that the standard definition of Harsanyi-Chwe. Hence our first result shows that, whenever a situation is a singleton stable set in the Harsanyi-Chwe sense, it is also a singleton stable set with heterogeneous expectations. However, we show through an example that a non-singleton Harsanyi-Chwe stable set need not be stable when agents have heterogeneous expectations. This observation is due to the fact that the Harsanyi-Chwe notion, contrary to ours, does not require agents to hold consistent expectations.

How does the relaxation of the hypothesis of common expectations expand the set of situations that can be supported as singleton farsighted stable sets? We explore this question by looking at two specific applications of farsighted stability.

In two-sided one-to-one matching problems, Mauleon, Vannetelbosch and Vergote [11] show that, with common expectations, a matching can be supported as a singleton farsighted stable set if and only if it is stable – individually rational and immune to blocking by pairs of agents. Allowing for heterogeneous expectations enables us to sustain a much larger set of matchings. Any matching which contains a "top matching" where all agents are matched to their top partners, and any matching that at least one agent strictly prefers to being matched to any agent for whom is the top partner can be supported as a singleton farsighted stable set.

In voting situations where the power structure is given by a simple game, with common expectations, an alternative is a singleton farsighted stable set if and only if it beats any other alternative – a generalization of the notion of Condorcet winner for simple majority. The existence of a Condorcet winner requires very strict conditions on preferences, so that singleton stable sets are unlikely to exist under common expectations. By contrast, when agents have heterogeneous expectations, the set of alternatives which can be supported is typically very large. To characterize it, we partition of the set of agents, putting together agents who have the same favorite alternative, $\pi = \{S_1, .., S_K\}$ with favorite alternatives $\{a_1, .., a_K\}$. If there is an agent i such that $N \setminus S_k \setminus i$ is a winning coalition for all k, then any alternative can be supported as a singleton farsighted stable set. Otherwise, we identify for every i those blocks of the partition S_k for which $N \setminus S_k \setminus i$ is not winning, and support any alternative that agent i prefers to all a_k.

The rest of the paper is organized as follows. In Section 2, we informally present two examples in which Harsanyi-Chwe farsighted stable sets do not exist and show how relaxing the hypothesis of common expectations allows us to support singleton farsighted stable sets. Section 3 is devoted to the presentation of the model and notations, and contains our preliminary results. Section 4 contains our results in the model of one-to-one matching and

Section 5 in the voting model with a simple game. We conclude in Section 6.

2 Two examples

In this Section, we consider two simple examples where the Harsanyi-Chwe farsighted stable set does not have any predictive power, and the introduction of heterogeneous expectations results in sharp predictions. First consider a roommate problem with three agents and cyclical preferences.[6] There is a single room which can accommodate two agents. There are three possible states corresponding to the three possible pairs: $a = \{1, 2\}$, $b = \{2, 3\}$ and $c = \{1, 3\}$. The consent of both players i and j is required to form the pair $\{i, j\}$ so that moves across states are shown by the following Figure:

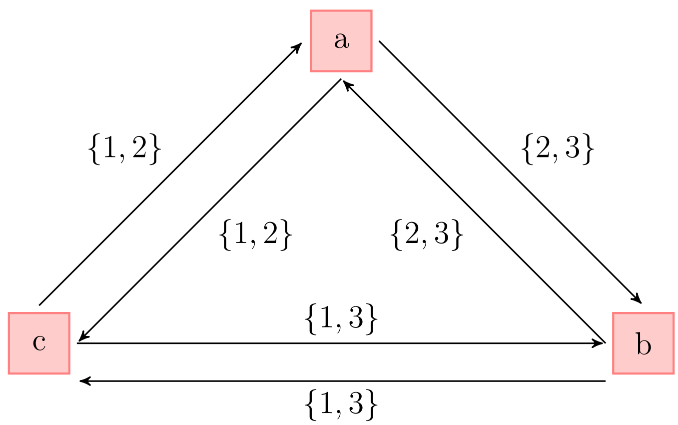

Figure 1: The roommate problem

Preferences are cyclical with $a \succ_1 c \succ_1 b$, $b \succ_2 a \succ_2 c$ and $c \succ_3 b \succ_3 a$. It is easy to check that when agents have common expectations, the following are the only (direct and indirect) dominance relations: b dominates a through a move of the coalition $\{2, 3\}$, c dominates b through a move of the coalition $\{1, 3\}$ and a dominates c through a move of the coalition $\{1, 2\}$. The core (the set of undominated states) is thus empty. The farsighted stable set is defined as a set of outcomes for which (i) no two elements of the set dominate each other (internal stability) and (ii) every element outside the set is dominated by an element in the set (external stability) . It is easy to check that the stable set is empty as internal stability can only be satisfied if the set is a singleton, and no state dominates the two other states.

[6]This Example also appears as Example 1 in Diamantoudi and Xue [3] and Example O.2 in the Supplemental Material of Ray and Vohra [14].

However, when agents have heterogeneous expectations, new dominance relations appear and it becomes possible to support any state as a singleton stable set.[7] Consider the following expectations: agent 1 believes that $b \to c \to a$ and that a is an absorbing set, agent 2 believes that $c \to a \to b$ and b is an absorbing set. Agent 3 believes that $a \to b \to c$ and c is an absorbing set. These expectations are path persistent. Furthermore, every agent believes that any move will lead to their favorite state, so every agent is willing to move along the paths: all moves are rationalizable and the expectations are consistent. Now from the point of view of agent 1, a dominates both b and c so $\{a\}$ is a singleton farsighted stable set. From the point of view of agent 2, b dominates both c and a and $\{b\}$ is a singleton farsighted stable set. From the point of view of agent 3, c dominates both a and b and $\{c\}$ is a singleton stable set. Hence, depending on the agent whose expectations are taken to anchor the dominance relation, any of the states can be supported as a singleton farsighted stable set.

The second example considers again three agents and three alternatives, a, b and c. Agents choose among alternatives through a simple majority vote. Hence every majority of two voters can implement a move from one alternative to another. Let $S_1 = \{1, 2\}, S_2 = \{2, 3\}$ and $S_3 = \{1, 3\}$ be the three two-player coalitions. We describe the situation through the following moves:

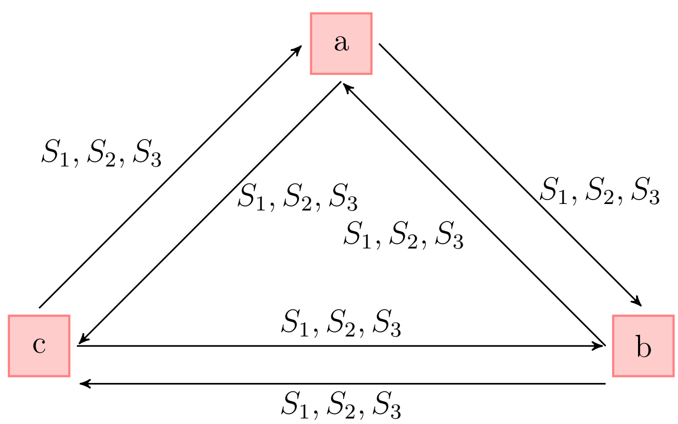

Figure 2: Majority voting

In this example, as opposed to the roommate problem, the same two-player coalitions are effective in any move from one state to another. Suppose that there exists a Condorcet cycle. Agents have cyclical preferences: $a \succ_1 b \succ_1 c$, $b \succ_2 c \succ_2 a$ and $c \succ_3 a \succ_3 b$. The only (direct and indirect) dominance relations are, as in the roommate example: a dominates b, b

[7] Diamantoudi and Xue [3] also support all pairs of agents fbut they use a very different argument. They suppose that agents have pessimistic expectations on the path following a move, so that no agent is willing to move out of a state where they get their second best outcome by fear of being left alone after the move.

dominates c and c dominates a, so that the core and farsighted stable sets are empty. When agents have heterogeneous expectations, as in the roommate example, we can support any alternative. Assume that each agent expects to attain their optimal alternative: 1 believes that $b \to c \to a$ and that a is an absorbing set, agent 2 believes that $c \to a \to b$ and b is an absorbing set. Agent 3 believes that $a \to b \to c$ and c is an absorbing set. Then $\{a\}$ is a farsighted stable set supported by the expectations of agent 1, $\{b\}$ is a farsighted stable set supported by the expectations of agent 2 and $\{c\}$ is a farsighted stable set supported by the expectations of agent 3.

These two examples show that, by relaxing the hypothesis of common expectations, we are able to construct nonempty stable farsighted sets and to support many situations (in fact all situations) as singleton farsighted stable sets. We will return to these two examples in Sections 4 and 5 where we fully characterize farsighted stable sets with common and heterogeneous expectations in one-to-one matching problems and voting models based on simple games.

3 Farsighted stability with heterogeneous expectations

In this Section, we define formally our new concept of farsighted stability. We first review the well-known notions of abstract systems, myopic and farsighted dominance and stable sets. We then introduce the concept of expectation functions, associating a path to each node for each agent. We define properties of path persistence, rationalizability and consistency of profiles of expectation functions. Finally, we use consistent expectation functions to define i-dominance and i-stable sets.

3.1 Abstract systems, stable sets, and farsighted stability

We consider an *abstract system* $\langle N, A, \{\succeq_i\}_{i \in N}, \{\to_S\}_{S \subseteq N,\ S \neq \emptyset}\rangle$ with a set of agents N and a set of alternatives A. Each agent $i \in N$ has preferences $\succeq_i$ over the alternatives in A. Coalitions $S \subseteq N,\ S \neq \emptyset$, of agents have the possibility to transition between alternatives according to the moves $\to_S$. For any two alternatives a and b, if $a \to_S b$, then S can move from alternative a to alternative b and we say that coalition S is effective for the transition from a to b.

In the abstract system $\langle N, A, \{\succeq_i\}_{i \in N}, \{\to_S\}_{S \subseteq N,\ S \neq \emptyset}\rangle$, alternative b is said to dominate alternative a via coalition S if $a \to_S b$ and $b \succ_i a$ for all $i \in S$. This is denoted by $b \text{ dom}_S\ a$. Alternative b is said to dominate alternative a, denoted by b dom a, if there exists a coalition S such that $b \text{ dom}_S\ a$.

Stable sets are defined with respect to dominance relations. A set of alternatives is a stable set when no two alternatives in the set dominate each other and each alternative that is not in the set is dominated by some alternative in the set.

Definition 3.1 $\rhd$-stable set *Given any dominance relation $\rhd$ on $A \times A$, a set $\Sigma \subseteq A$ of alternatives is a $\rhd$-stable set if it satisfies the following two conditions.*

1. *Internal stability: There do not exist two alternatives $a, b \in \Sigma$ such that $a \rhd b$.*
2. *External stability: For any alternative $b \in A \setminus \Sigma$ there exists an alternative $a \in \Sigma$ such that $a \rhd b$.*

In the abstract system derived from a characteristic function game (N, V), in which V is a correspondence that assigns a non-empty set of feasible utility vectors $V(S) \subseteq \Re^S$ to each coalition of agents $S \subseteq N$, the set of alternatives is $V(N)$, for two alternatives $x, y \in V(N)$ and coalition S it holds that $x \to_S y$ iff $y_{|S} \in V(S)$, and agent i's preferences are given by $x \succeq_i y$ iff $x_i \geq y_i$. The stable sets with respect to the resulting domination relations x dom y are known as the von Neumann-Morgenstern stable sets.

Definition 3.2 Von Neumann-Morgenstern stable set *A von Neumann-Morgenstern stable set in a characteristic function game (N, V) is a set $\Sigma \subseteq V(N)$ such that for any two $x, y \in \Sigma$ there is no $S \subseteq N$ with x dom$_S$ y, and for any $y \in V(N) \setminus \Sigma$ there exists an $x \in \Sigma$ and an $S \subseteq N$ with x dom$_S$ y.*

In a von Neumann-Morgenstern stable set, players and coalitions consider one-step, myopic moves. Letting agents consider sequences of moves leads to the definition of farsighted dominance. Formally, in an abstract system $\langle N, A, \{\succeq_i\}_{i \in N}, \{\to_S\}_{S \subseteq N,\ S \neq \emptyset}\rangle$, farsighted dominance is defined as follows.

Definition 3.3 Farsighted dominance *Alternative b is said to farsightedly dominate alternative a if there exists a sequence of alternatives $a = a_0, a_1, \ldots, a_k = b$ such that for each $j \in \{1, 2, \ldots, k\}$ there exists a coalition S_j of agents such that $a_{j-1} \to_{S_j} a_j$ and $b \succ_i a_{j-1}$ for all $i \in S_j$.*

Thus, a is farsightedly dominated by b if there is a sequence of coalitions who can move from alternative a to alternative b and such that each agent prefers the final alternative b to any intermediate alternative for which the agent's participation is needed to transition out of that alternative on the path to b.

Stable sets with respect to the farsighted dominance relation are called farsighted stable sets.

Definition 3.4 Farsighted stable set *A a set $\Sigma \subseteq A$ of alternatives is a farsighted stable set if there are no two alternatives $a, b \in \Sigma$ such that b farsightedly dominates a, and for any $a \in A \setminus \Sigma$ there exists a $b \in \Sigma$ that farsighted dominates a.*

3.2 Expectation Functions

Farsighted dominance as defined in the previous subsection requires coordination of the expectations of the agents. An alternative a is farsightedly dominated if and only if there exists an alternative b and sequence of moves from a to b such that all effective agents anticipate that b will be reached and prefer alternative b to the (intermediate) alternative from which they are moving. We want to allow for agents to hold different expectations. To that end, we focus attention on players' expectations over the sequences of moves and define explicitly expectation functions, using paths of alternatives that can be supported by coalitional moves.

Definition 3.5 Path *A path from alternative a to alternative b is a sequence of different alternatives $(a_1, a_2, a_3, \ldots, a_{k+1})$ such that $a_1 = a$, $a_{k+1} = b$, and for each $j = 1, 2, \ldots, k$ there exists a coalition S_j such that $a_j \to_{S_j} a_{j+1}$. The path is said to originate from a. If $k = 0$, then the path equals (a).*

Each agent i has expectations on the paths that will be followed in the abstract system. An expectation function associates to any alternative a the unique path of moves that player i expects from a

Definition 3.6 Expectation function *For each agent $i \in N$, the expectation function of i, p_i, assigns to every alternative $a \in A$ a path $p_i(a)$ originating from a.*

Agent i expects that if the system gets to alternative a, then the path $p_i(a)$ will be followed from there on. If the agent expects the system to stay in alternative a, then $p_i(a) = (a)$. Notice that this definition requires the players to expect a single deterministic sequence of moves to be followed at each alternative. We require the expectations of an agent to be persistent in the following sense.

Definition 3.7 Path persistence *The expectation function p_i is path persistent if for every $a \in A$, if $p_i(a) = (a_1, a_2, a_3, \ldots, a_{k+1})$, then for every $j = 1, 2, \ldots, k$ it holds that $p_i(a_j) = (a_j, a_{j+1}, \ldots, a_{k+1})$.*

Path persistence requires that when an agent expects to follow a path from an alternative a to an alternative b via an intermediate alternative $a_j \neq a, b$, then the agent expects the continuation of that same path from alternative a_j. This condition requires that a player sticks to his expectations throughout the path originating at any node. It rules out expectations where the agent, once at the intermediate alternative a_j, either expects to not continue (i.e., $p_i(a_j) = (a_j)$), or to continue in a different direction (i.e., the first step on the path $p_i(a_j)$ is to an alternative different from a_{j+1}), or to continue in the same direction, but to end up at a different alternative eventually. Path persistence of an expectation function also implies a form of path independence because an agent will expect the first step on the path $p_i(a_j)$ to be to alternative a_{j+1} regardless of how they expected to arrive at alternative a_j. Note that path dependence rules out that an agent always looks ahead k steps and thus looks ahead one step further once the first step on a path is taken.[8]

While path persistence relates agent i's expectations from various alternatives to each other, it does not imply any restrictions on relations between expectations of different agents. However, the cooperation of other agents may be needed to move between alternatives. These agents will only cooperate if they expect to end up with an alternative that is better for them. The formal definition the resulting consistency condition requires us to introduce some notation.

Let $\mathbf{p} = \{p_i\}_{i \in N}$ denote a set of expectation functions, one for each agent. For each player i and each alternative a, define $t(p_i, a)$ to be the terminal node of the path $p_i(a)$: If $p_i(a) = (a, a_2, a_3, \ldots, a_{k+1})$, then $t(p_i, a) = a_{k+1}$. Note that path persistence implies that $t(p_i, a_j) = t(p_i, a))$ for all alternatives a_j on the path $p_i(a) = (a, a_2, a_3, \ldots, a_{k+1})$.

Definition 3.8 Rationalizable transitions *A transition from an alternative a to an alternative b is rationalizable by $\mathbf{p}$ if there exists an $S \subseteq N$ such that $a \to_S b$ and for all $i \in S$ it holds that $p_i(a) = (a, b, \ldots)$ and $t(p_i, a) \succ_i a$.*

Rationalizability of a transition by a set of expectation functions has two elements. The first element is effectivity: It must be that a coalition S is effective for the transition and that all the agents $i \in S$ expect to make the transition as a first step on the path $p_i(a)$. The second element is that all players $i \in S$ agree to make the transition from a to b because they expect to end up at an alternative better than a. Notice that we require all players in S to strictly prefer the alternative $t(p_i, a)$ to the alternative a. If some players are indifferent, we assume that they will not move out of the status quo.

We restrict the expectations of the agents by requiring that agents can only expect rationalizable transitions and we call such expectations consistent.

Definition 3.9 Consistent expectations *The set of expectations $\mathbf{p} = \{p_i\}_{i \in N}$ is consistent if for every agent i and every alternative a, every step on the path $p_i(a)$ is rationalizable by $\mathbf{p}$.*

3.3 Farsighted stability with heterogeneous expectations

We now construct a dominance relation based on consistent expectations. This construction is not obvious. When players hold different expectations, an alternative may be dominated

[8]This condition thus rules out k-farsightedness as proposed by Herings, Mauleon and Vannetelbosch [7]. We believe that k-farsightedness poses important conceptual problems: it is not clear how a player should anticipate what will happen beyond his fixed horizon, and the definition of sequences of moves of length greater than k becomes problematic.

(in the sense that there is a sequence of rationalizable moves leading out of it) or not depending on the agents' point of view. As there is no simple way out of this problem, we adopt a very permissive concept of dominance. We say that an alternative b dominates an alternative a through agent i if agent i expects a path from a ending at b. Formally, we define:

Definition 3.10 ***i*-Dominance** *Given a consistent set of expectations* $\mathbf{p} = \{p_i\}_{i\in N}$ *and an agent* $i \in N$*, alternative* b *i-dominates alternative* a*, denoted* $b \rhd_{i,\mathbf{p}} a$*, if* $b = t(p_i, a)$ *and* $b \neq a$*.*

We are interested in stable sets with respect to these dominance relations. Keeping in mind that we consider the stability through the eyes of a specific agent i for a fixed profile of expectation functions $\mathbf{p}$, we define $\mathbf{p}$-farsighted i-stable sets as follows.

Definition 3.11 **p-Farsighted *i*-stable set** *Given a consistent set of expectations* $\mathbf{p} = \{p_i\}_{i\in N}$ *and an agent* $i \in N$*, a* $\mathbf{p}$*-farsighted i-stable set is a set of alternatives* $\Sigma_{i,\mathbf{p}}$ *that satisfies internal and external stability*[9] *with respect to dominance relation* $\rhd_{i,\mathbf{p}}$*.*

Farsighted stable sets with heterogeneous expectations are those sets of alternatives that can supported by appropriate expectations.

Definition 3.12 **Farsighted stable set with heterogeneous expectations** *A farsighted stable set with heterogeneous expectations is a set of alternatives* Σ *such that there exists a consistent set of expectation functions* $\mathbf{p} = \{p_i\}_{i\in N}$ *and an agent* $i \in N$ *that support* Σ *as a* $\rhd_{i,\mathbf{p}}$*-stable set.*

Different consistent expectation functions give rise to different farsighted stable sets with heterogenous expectations. Notice that the expectation functions where all agents expect to remain at any alternative trivially satisfy path-persistence and consistency. These expectation functions support the entire set of alternatives, A, as a stable set. Hence, as opposed to classical notions of farsighted stable sets when agents hold homogeneous expectations, the existence of a farsighted stable set with heterogeneous expectations is guaranteed.

Proposition 3.13 **Existence of farsighted stable sets with heterogeneous expectations** *The set* A *of all alternatives is a farsighted stable set with heterogeneous expectations in the abstract system* $\langle N, A, \{\succeq_i\}_{i\in N}, \{\to_S\}_{S\subseteq N,\ S\neq\emptyset}\rangle$*.*

Proof: Considering the expectation functions $\{p_i\}_{i\in N}$ with $p_i(a) = (a)$ for all $a \in A$ and $i \in N$. These simple expectation functions reflect that no player expects any transitions between alternatives and they vacuously satisfy path persistence as well as rationalizability

[9] See definition 3.1

of all expected transitions. With these consistent expectations, no alternative i-dominates any other alternative for any agent $i \in N$. Thus, the set of all alternatives is $\rhd_{i,\mathbf{p}}$-stable for all i. ■

The previous theorem establishes a marked difference between farsighted stable sets with heterogeneous expectations and farsighted stable sets as in Definition 3.4. While farsighted stable sets with homogeneous expectations may not exist, farsighted stable sets with heterogeneous expectations always exist. We also note that if an alternative b dominates an alternative a when agents hold homogeneous expectations, it will also dominate the alternative a when agents hold heterogeneous expectations. Hence, if a singleton $\{a\}$ is a farsighted stable set with homogeneous expectations, it is also a farsighted stable set with heterogeneous expectations.

Proposition 3.14 A singleton farsighted stable set with heterogeneous expectations *Let $\{a\}$ be a singleton farsighted stable set with homogeneous expectations in the abstract system $\langle N, A, \{\succeq_i\}_{i\in N}, \{\rightarrow_S\}_{S\subseteq N,\ S\neq\emptyset}\rangle$. Then $\{a\}$ is also a farsighted stable set with heterogeneous expectations.*

Proof: We define a consistent set of expectation functions $\mathbf{p} = \{p_i\}_{i\in N}$ and an agent $i \in N$ that support $\{a\}$ as a $\rhd_{i,\mathbf{p}}$-stable set. The definition is recursive.

Initial step. Pick an alternative $b \in A \setminus \{a\}$. Because of external stability of the farsighted stable set with homogeneous expectations $\{a\}$, we know that we can find a $k \geq 1$ and a sequence of alternatives $b = b_0, b_1, \ldots, b_k = a$ such that for each $j \in \{1, 2, \ldots, k\}$ there exists a coalition S_j of agents such that $b_{j-1} \rightarrow_{S_j} b_j$ and $a \succ_i b_{j-1}$ for all $i \in S_j$.[10] For each $i \in N$ and each $j \in \{0, 1, \ldots, k\}$, define $p_i(b_j) := (b_j, b_{j+1}, \ldots, b_k = a)$. We have now defined $p_i(\tilde{a})$ for all $i \in N$ and $\tilde{a} \in A_1 := \{b_0, b_1, \ldots, b_k\}$. Note that $\{a, b\} \subseteq A_1$. Also note that the (possibly partial) expectation functions that we have defined so far satisfy path persistence and that they rationalize all transitions on these paths. Moreover, $t(p_i, \tilde{a}) = a$ for all expectations so far defined.

Induction step. Suppose that we have defined a set of alternatives $A_m \subseteq A$ with $a \in A_m$, and expectations $p_i(\tilde{a})$ for all $i \in N$ and $\tilde{a} \in A_m$. If $A_m = A$, then we are done defining the expectation functions $\{p_i\}_{i\in N}$. Otherwise, pick an alternative $c \in A \setminus A_m$. Because of external stability of the farsighted stable set with homogeneous expectations $\{a\}$, we know that we can find a $l \geq 1$ and a sequence of alternatives $c = c_0, c_1, \ldots, c_l = a$ such that for each $j \in \{1, 2, \ldots, l\}$ there exists a coalition T_j of agents such that $c_{j-1} \rightarrow_{T_j} c_j$ and $a \succ_i c_{j-1}$ for all $i \in T_j$. If none of the intermediate alternatives $c_1, c_2, \ldots, c_{l-1}$ are in A_m, then we can just define $p_i(c_j) = (c_j, c_{j+1}, \ldots, c_l = a)$ for each $i \in N$ and each $j \in \{0, 1, \ldots, l-1\}$. However, to make sure that we maintain path persistence, we need to be a bit more careful in case $\{c_1, c_2, \ldots, c_{l-1}\} \cap A_m \neq \emptyset$. We cover both possibilities by defining $\tilde{l}$ such that $c_{\tilde{l}} \in A_m$ and $c_j \notin A_m$ for each $j < \tilde{l}$. Because $c_0 = c \in A \setminus A_m$ and $c_l = a \in A_m$, we know that

[10] See Definition 3.3.

$1 \leq \tilde{l} \leq m$. We define $A_{m+1} := A_m \cup \{c_0, \ldots, c_{\tilde{l}-1}\}$. Note that $A_{m+1} \supset A_m$. For each $\tilde{a} \in \{c_0, \ldots, c_{\tilde{l}-1}\}$, the expectations $p_i(\tilde{a})$ are defined by piecing together the new path until an existing one is reached and then continuing on that from there on: For each $i \in N$ and $j \in \{0, \ldots, \tilde{l}-1\}$, define $p_i(c_j) := (c_j, c_{j+1}, \ldots, c_{\tilde{l}-1}, p_i(c_{\tilde{l}}))$. Note that all expectations $p_i(\tilde{a})$ so far defined have a as the terminal node, $t(p_i, \tilde{a}) = a$, and they satisfy path persistence and rationalize all transitions on these paths.

Because expectations are defined for additional alternatives in every step, the procedure described above results in the definition of complete expectation functions $\mathbf{p} = \{p_i\}_{i \in N}$. By construction, these expectation functions satisfy path persistence and are consistent. Moreover, $t(p_i, \tilde{a}) = a$ for each $i \in N$ and $\tilde{a} \in A$. Thus, $\{a\}$ satisfies $\triangleright_{i,\mathbf{p}}$-external stability. Internal stability is of course immediate for any singleton-set. We conclude that $\{a\}$ is a $\triangleright_{i,\mathbf{p}}$-stable set for any $i \in N$ and a farsighted stable set with heterogeneous expectations. ∎

While a singleton farsighted stable set with homogeneous expectations is always a farsighted stable set with heterogeneous expectations, a similar implication is not true for farsighted stable sets consisting of more than one alternative. We demonstrate this in the following example, which is taken from Dutta and Vohra [4].

Example 3.15 *Consider an* abstract system $\langle N, A, \{\succeq_i\}_{i \in N}, \{\rightarrow_S\}_{S \subseteq N,\ S \neq \emptyset}\rangle$ *with three agents, five alternatives, and possible transitions as depicted in Figure 3. The preferences of the agents are transitive and given by* $e \succ_1 a \sim_1 b \succ_1 c \sim_1 d$, $d \succ_2 a \sim_2 b \succ_2 c \sim_2 e$, *and* $d \sim_3 e \succ_3 a \sim_3 b \succ_3 c$.

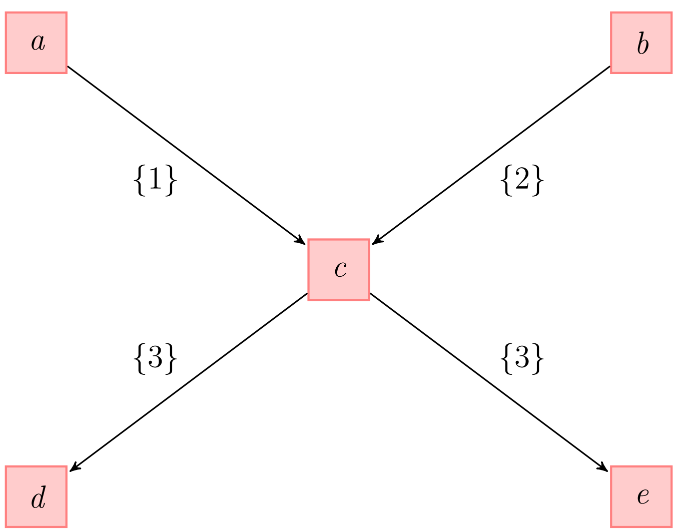

Figure 3: The Harsanyi-Chwe farsighted stable set is not stable with heterogeneous expectations

In this abstract system, there exists a unique farsighted stable set with homogeneous expectations and it is $\{d, e\}$. *Alternative* e *farsightedly dominates alternatives* a *and* c *through*

$a \rightarrow_{\{1\}} c \rightarrow_{\{3\}} e$ and alternative d farsightedly dominates alternatives b and c through $b \rightarrow_{\{2\}} c \rightarrow_{\{3\}} d$. The paths that are used in these farsighted domination relations do not satisfy path persistence because they depend on agent 3 transitioning from alternative c sometimes to alternative d and sometimes to alternative e.

For any consistent set of expectation functions $\mathbf{p}$, we need $p_3(c) = (c, d)$ to have the possibility of alternative d $\rhd_{i,\mathbf{p}}$-dominating any other alternative (because $c \rightarrow_3 d$ needs to be rationalizable by $\mathbf{p}$). However, if $p_3(c) = (c, d)$, then the only possible $\mathbf{p}$-rationalizable expectations of agent 1 at alternative a are (a), (a, c), and (a, c, d). The latter two expectations, however, are ruled out because $a \succ_1 c \sim_1 d$. Thus, $p_1(a) = a$ has to hold and alternative a is not $\rhd_{i,\mathbf{p}}$-dominated for any agent i. This demonstrates that $\{d, e\}$ is not a farsighted stable set with heterogeneous expectations, because it fails the external stability criterion.

Proposition 3.13 demonstrates that any set B of alternatives that satisfies internal stability with respect to a dominance relation $\rhd_{i,\mathbf{p}}$ can be expanded to one that also satisfies external stability with respect to $\rhd_{i,\mathbf{p}}$ by defining any agent j's expectations at alternatives $c \in A \setminus B$ that are not $\rhd_{i,\mathbf{p}}$-dominated by some alternative in B to equal $p_j(c) = (c)$ and adding these alternatives c to the set B. Example 3.15 demonstrates that the introduction of path persistent expectation functions eliminates the possibility that the same agent holds different expectations at the same alternative. This reduces the possibility of domination and may prevent some non-singleton sets of alternative to emerge as stable sets when agents are characterized by expectation functions. Given these considerations, we focus our attention on singleton farsighted stable sets with heterogeneous expectations.

4 One-to-one matching

We consider a general model of one-to-one matching which encompasses both two-sided and one-sided matching. This is a special case of the model of hedonic coalitions of Diamantoudi and Xue [3] but it generalizes the one-to-one two-sided matching model of Mauleon, Vannetelbosch and Vergote [11]. There is a collection $\mathcal{F}$ of feasible *pairs* that players can form.[11] A *matching* μ is a one-to-one mapping from N to itself such that, (i) $\mu(\mu(i)) = i$ and (ii) if $\mu(i) \neq i$, $\{i, \mu(i)\} \in \mathcal{F}$. The set of all possible matchings is denoted $\mathcal{M}$. If $\mu(i) = i$ agent i remains single, and if $\mu(i) \neq i$ agent i has a partner $\mu(i)$.

For any player i, let $\mathcal{F}(i) = \{j \in N \mid \{i, j\} \in \mathcal{F}\}$ denote the set of potential partners of agent i. Every agent i has a strict preference $\succ_i$ over the set $\mathcal{F}(i) \cup \{i\}$, expressing his ranking of potential partners. We suppose that for every agent i there exists a potential partner that they prefer to remaining single: For any i, there exists a $j \neq i$ such that $\{i, j\} \in \mathcal{F}$ and $j \succ_i i$. We let $\succeq_i$ denote the weak preference corresponding to $\succ_i$.

[11] In two-sided matching problems, the set N is partitioned into a set of men and women and a feasible pair must contain a man and a woman; in one-sided matching problems, any pair of agents is feasible.

A matching μ is *individually rational* if and only if $\mu(i) \succ_i i$ for all $i \in N$. A matching μ is *blocked* by a pair of agents $\{i,j\}$ if and only if $j \succ_i \mu(i)$ and $i \succ_j \mu(j)$.

Definition 4.1 Stable matching *A matching is* stable *if it is individually rational and cannot be blocked by any pair $\{i,j\}$.*

As is well-known, existence of a stable matching is guaranteed in two-sided matching problems but not when the matching problem is one-sided.

In order to consider farsighted stable sets with heterogeneous expectations in the matching setting, we need to extend the preferences of agents from preferences over partners to preferences over matchings. We do this in a natural way by supposing that agents do not care about matches formed by other agents, so that for any two matchings μ and μ' it holds that $\mu \succ_i \mu'$ if and only if $\mu(i) \succ_i \mu'(i)$. In particular, agent i is indifferent among all matchings μ and μ' under which he is matched to the same partner. We also extend the definition of blocking to coalitions of players of arbitrary size.

Definition 4.2 Blocking by coalitions *A matching μ is* blocked *by a coalition S if there exists another matching μ' such that $\mu'(S) = S$ and $\mu' \succ_i \mu$ for all $i \in S$.*

In words, a coalition S can block a matching μ through a matching μ' if (i) in μ' all agents in S are matched to other agents in S and (ii) all agents in S strictly prefer their partner under μ' to their partner under μ. In one-to-one matching problems, effective coalitions have size one and two. Hence larger coalitions will only block by forming pairs and singletons. We conclude that a matching is not blocked by any coalition if and only if it is stable.

We complete the description of the abstract system corresponding to one-to-one matching problems by specifying the effectivity function. Every agent can unilaterally leave a partner that they are currently matched with, but forming a new partnership requires the consent of both agents. This yields the following definition of effectivity.[12]

Definition 4.3 Effectivity *A coalition S is* effective *in the move from a matching μ to a matching μ' if the following two conditions hold: 1. $\mu'(i) \notin \{i, \mu(i)\}$ implies $\{i, \mu'(i)\} \subseteq S$, and 2. $\mu'(i) = i \neq \mu(i)$ implies $\{i, \mu(i)\} \cap S \neq \emptyset$.*

We first recall the results obtained by Diamantoudi and Xue [3] and Mauleon, Vannetelbosch and Vergote [11] on farsighted stability in hedonic coalitions and one-to-one two-sided matchings. Diamantoudi and Xue [3] show that any partition which does not belong to

[12] This definition of effectivity is the definition given by Mauleon, Vannetelbosch and Vergote [11] and it can be traced back to a condition in Roth and Sotomayor [15]. It differs from the effectivity function used by Ehlers [5] , who assumes that a coalition S is effective in the move from a matching μ to a matching μ' whenever $\mu'(S) = \mu(S)$, implicitly assuming that a coalition S can force agents not in S to match with different agents.

the core of the hedonic game is indirectly dominated by a partition in the core.[13] They construct a sequence of moves, where players initially break coalitions into singletons and then rebuild coalitions to reach the core-stable partition. This construction however does not help establish the existence (or the absence) of indirect domination paths between core-stable partitions. Diamantoudi and Xue [3] obtain a stronger result when the hedonic game satisfies the top-coalition property of Banerjee, Konishi and Sönmez (2001).

Definition 4.4 Top coalition *A coalition I is a top coalition in N if all players in I strictly prefer I to any other coalition.*

Banerjee, Konishi and Sönmez iteratively define a top-coalition partition by first identifying a top coalition I_1 in N, then a top coalition I_2 in $N \setminus I_1$, and so on. When a top-coalition partition π^* exists, it is the unique core-stable partition in the hedonic game. Diamantoudi and Xue [3] then show that the singleton $\{\pi^*\}$ is the unique farsighted stable set in the hedonic game.

Mauleon, Vannetelbosch and Vergote [11] specialize the model to one-to-one two-sided matchings and obtain a stronger result: They prove existence of indirect domination paths from any matching (including stable matchings) to any stable matching. This shows that for any stable matching μ, the singleton set $\{\mu\}$ is a farsighted stable set. Furthermore, they prove that there cannot be indirect domination paths from any matching to a matching that is not stable. Hence, any stable matching is a singleton farsighted stable set, and any farsighted stable set is a singleton consisting of a stable matching. When a top-coalition matching μ^* exists, then $\{\mu^*\}$ is the unique farsighted stable set.

Both the results of Diamantoudi and Xue [3] and Mauleon, Vannetelbosch and Vergote [11] apply to situations where the core is nonempty. If the core is empty (as in the cyclical roommate problem presented in Section 2), their characterizations cannot be used and the farsighted stable set may not exist.

We now turn to the model of farsighted stability with heterogeneous expectations. In order to simplify the arguments, we will assume that all agents are acceptable partners and that the worst outcome for any agent is to remain single:

Assumption 4.5 *For any i any $j \neq i$, $j \succ_i i$.*

We define top-match matchings in the one-to-one matching model. For any set of agents S, let $\nu_S(i)$ be the top partner of agent i in the set S and $\nu(i) = \nu_N(i)$ – the top partner of agent i in the entire set N. Given Assumption 4.5, for each $S \subseteq N$ with at least two agents ($|S| > 1$) it holds that $\nu_S(i) \succ_i i$ for each $i \in S$.

[13] They assume the following effectivity function: A coalition S is effective in the move from a partition π to a partition π' if all players in S form a block in π' and the players who belonged to blocks T in π such that $T \cap S \neq \emptyset$ stick together and form the blocks $T \setminus S$ in π'.

Definition 4.6 Top match *For any two agents $i, j \in N$ such that $\nu(i) = j$ and $\nu(j) = i$, agents i and j are each other's top partners and $\{i, j\}$ a top match in the set N.*

Analogously to a top-coalition partition, a top-match matching is defined recursively: In the initial step, let $S_1 = N$ and $I_1 = \{i \in N \mid \nu(\nu(i)) = i\}$, the set of agents who are part of a top match in N. For all $i \in I_1$, define $\mu^*_{I_1}(i) = \nu(i)$, so that $\mu^*_{I_1}$ identifies all the top matches in N. We now remove all top-matched agents from the set N and let $S_2 = N \setminus I_1$ and identify the agents who are part of a top match in S_2. Formally, having defined $S_1, \ldots, S_{k-1}$ and $I_1, \ldots, I_{k-1}$, in the recursion step we define $S_k = S_{k-1} \setminus I_{k-1}$ and $I_k = \{i \in S_k \mid \nu_{S_k}(\nu_{S_k}(i)) = i\}$, the set of agents who are part of a top match in S_k. For all $i \in I_k$, define $\mu^*_{I_k}(i) = \nu_{S_k}(i)$, so that $\mu^*_{I_k}$ identifies all the top matches in S_k. Note that in the set $S_k \subset N$ it may be the case that $\nu_{S_k}(i) = i$ for some agent $i \in S_k$, in which case agent i is matched to themselves and $\mu^*_{I_k}(i) = i$.

Because the set of agents is finite, the recursion will end in finite time either because all agents have been matched or because there are no top matches among the agents that are left. Let K be the final step of the recursion, so that either $S_K = \emptyset$ or $I_K = \emptyset$. We let $I = \cup_{i=1}^{K-1} I_k$ denote the set of agents who are recursively matched to their top partners, and $J = N \setminus I$ the remaining agents (i.e., those in S_K). Note that for any $j \in J$ it holds that $\nu_J(j) \neq j$ (and thus if $J \neq \emptyset$, then $|J| \geq 2$). Also note that I is empty if there are no top matches in the set N. If I is nonempty, the we have defined the unique top-match matching μ^*_I of agents in I ($\mu^*_I(i) = \mu^*_{I_k}$ for $i \in I_k$), and possibly $\mu^*_I(i) = i$ for some $i \in I$.

Our first result is a preliminary lemma that identifies particular cycles in sets of agents in which top matches do not exist.

Lemma 4.7 *Consider a set J of agents such that $\nu_J(\nu_J(i)) \neq i$ for every $i \in J$. Then it is possible to assign the numbers $1, 2, \ldots, M$, where $M = |J|$, to the agents in J in such a way that $\nu_J(m+1) \neq m$ for all $m = 1, .., M-1$, and $\nu_J(1) \neq M$.*[14]

Proof: Construct a directed graph in which the nodes are the agents in J and for any two agents i and J, $i \to j$ if and only if $j = \nu_J(i)$. Because $\nu_J(\nu_J(i)) \neq i$ for every $i \in J$, any directed cycles in the graph are of length greater or equal to 3. Fix an agent 1, and construct the sequence $2 = \nu_J(1), 3 = \nu_J(2), \ldots$. Eventually, we will come across an agent $k \geq 3$ such that $\nu_J(k) \in \{1, 2..., k-2\}$. For all $m \leq k-1$ the condition $\nu_J(m+1) \neq m$ is always satisfied because $m + 1 = \nu_J(m)$ and $\nu_J(\nu_J(m)) \neq m$.

If $k = M$, then $\nu_J(1) \neq M$ because $M \neq 2$. Thus, we have identified a cycle among all the agents with the desired properties.

If $k < M$, we distinguish between two cases. **Case 1.** If there exists an agent l in $J \setminus \{1, ...k\}$ such that $\nu_J(l) \neq k$, then let $k + 1 = l$, and continue, as before, with $k + 2 =$

[14] $|J|$ denotes the number of elements in J.

$\nu_J(k+1)$, $k+3 = \nu_J(k+2), \ldots$. For the same reasons as before, eventually, we will come across an agent $\tilde{k} \geq k+1$ such that $\nu_J(\tilde{k}) \in \{1, 2..., \tilde{k}-2\}$. For all $m \leq \tilde{k}-1$ the condition $\nu_J(m+1) \neq m$ is satisfied by construction. If necessary, we repeat the algorithm described in this case until we either have numbered all agents in J (as described above) or end up in Case 2. **Case 2.** If $\nu_J(l) = k$ for all agents $l \in J \setminus \{1, ...k\}$, then order the $M-k$ agents in $J \setminus \{1, ...k\}$ in some arbitrary manner and write $J \setminus \{1, ...k\} = \{l_1, l_2, \ldots, l_{M-k}\}$. The sequence $1, 2, \ldots, k-2, k-1, l_1, l_2, \ldots, l_{M-k}, k$ encompasses all agents and has the desired properties, because $\nu_J(l) = k$ for all $l \in \{l_1, l_2, \ldots, l_{M-k}\}$, $\nu_J(k) \notin \{l_1, l_2, \ldots, l_{M-k}\}$, and $\nu_J(1) \neq k$ because $k \neq 2$. ■

Lemma 4.7 shows that, in a set without any top matches, we can construct a cycle of agents $1, 2.., M, 1$ such that no agent in the cycle is the top partner of the succeeding agent in the cycle. This technical result will be very useful to construct chains of indirect dominance.

The analysis of farsighted stability with heterogeneous expectations will be separated into two cases, one where $I \neq \emptyset$ and one where $I = \emptyset$. We start with the first situation, when top matches exist in N.

Proposition 4.8 *Suppose that $I \neq \emptyset$. Then for any matching μ it holds that $\{\mu\}$ is a singleton farsighted stable set with heterogeneous expectations if and only if $\mu(i) = \mu_I^*(i)$ for each $i \in I$.*

Proof: (Sufficiency) Let μ_J be an arbitrary matching of the agents in J. We prove that the matching $\mu = (\mu_I^*, \mu_J)$ of agents in N forms a singleton farsighted stable set with heterogeneous expectations. We construct consistent expectation functions such that from any matching $\mu' \neq \mu$, all agents in I expect to follow a path ending at μ. Because of path persistence and consistency, we can represent the expectation functions using a graph with transitions between matchings. Figure 4 illustrates the construction of the expectation functions that rationalize the move from μ' to μ.

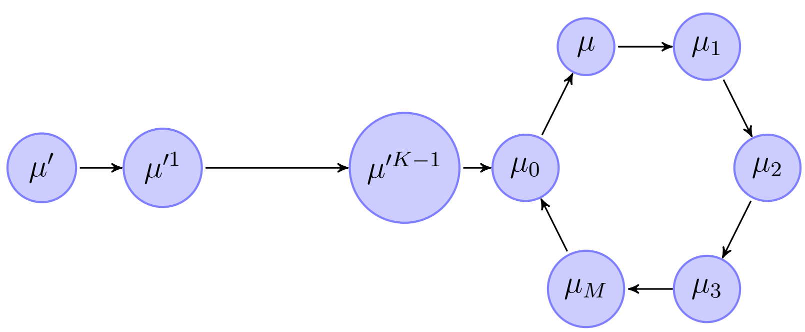

Figure 4: Expectation functions when $I \neq \emptyset$

To simplify notation in what follows, we introduce for any matching $\tilde{\mu}$ of agents in a set T and for any set of agents $S \subseteq T$ the matching $\tilde{\mu}_S$ of agents in S as follows: $\tilde{\mu}_S(i) = i$ for all agents $i \in S$ who were in $\tilde{\mu}$ matched to an agent in $T \setminus S$, and $\tilde{\mu}_S(i) = \tilde{\mu}(i)$ for all agents $i \in S$ such that $\tilde{\mu}(i) \in S$.

If $J \neq \emptyset$, then we have to find a way to make the agents in J form the matches in μ_J among themselves and to make the transitions between matchings rationalizable. In what follows, all the agents in I are matched to each other according to μ_I^*. We define the matching μ_0 by $\mu_0(i) = \mu_I^*(i)$ for all $i \in I$ and $\mu_0(j) = j$ for all $j \in J$, i.e.,

$$\mu_0 = (\mu_I^*, \{j\}_{j \in J}).$$

For each $j \in J$, we define the matching

$$\mu_j = (\mu_I^*, \{j, \nu_J(j)\}, \{k\}_{k \in J,\, k \neq j, \nu_J(j)}).$$

Thus, $\mu_j(i) = \mu_I^*(i)$ for all $i \in I$ and the match $\{j, \nu_J(j)\}$ between agent j and their preferred partner in J is the unique match among the agents in J. All other agents in J are single. Because $\nu_J(j) \neq j$ for all $j \in J$, all the matchings μ_j are clearly different from μ_0 and because $\nu_J(\nu_J(j)) \neq j$ for all agents in J (by definition of J), all these matchings are also different from one another. Note that it is possible that for the matching $\mu = (\mu_I^*, \mu_J)$, it holds that $\mu_0 = \mu$ or $\mu_j = \mu$ for some $j \in J$.

Using Lemma 4.7, we assign the numbers $1, 2, \ldots, M$, where $M = |J|$ to the agents in J in such a way that $\nu_J(m+1) \neq m$ for all $m = 1, 2, \ldots, M-1$, and $\nu_J(1) \neq M$. This defines a directed cycle $\mu_0, \mu, \mu_1, \ldots, \mu_M, \mu_0$. If $\mu_j = \mu$ for some $j \in J$, then re-number so that that particular agent j gets the number 1. In this case, the two nodes for μ_1 and μ are superimposed in Figure 4. If it is the case that $\mu_0 = \mu$, then the two nodes for μ_0 and μ are superimposed in Figure 4.

We give the agent in $j \in J$ the path persistent expectations represented by the path $\mu_{j+1} \to \mu_{j+2} \to \ldots \to \mu_0 \to \mu \to \mu_1 \to \ldots \to \mu_j$. Thus, from any matching along the cycle they expect a path that follows the cycle and terminates at μ_j.

We check that all the transitions in these path persistent expectations are rationalizable:

- The transition from μ_0 to μ (only if $\mu_0 \neq \mu$): Only players in J can be effective in this transition. Every agent $j \in J$ is willing to cooperate in the transition from μ_0 to μ because they expect to end up in the matching μ_j in which they are matched to their preferred partner in J, and they strictly prefer that match to being single.

- The transition from μ to μ_1 (only if $\mu \neq \mu_1$): Only players in J can be effective in this transition. For every pair of agents $j, k \in J$ that are matched in μ, it holds that either $j \neq \nu_J(k)$ or $k \neq \nu_J(j)$ (or both). If $j \neq \nu_J(k)$, then agent k is willing to cooperate in the transition from μ to μ_1 because they expect to end up in the matching μ_k in which they are matched to their preferred partner $\nu_J(k)$ in J. Thus, from each pair of agents that are matched in μ, we can select an agent who is willing to break their match and cooperate in the transition from μ to μ_1. It remains to show that the two agents 1 and $\nu_J(1)$ are willing to form the match $\{1, \nu_J(1)\}$. Agent 1 is willing to do this because they expect to end up in the matching μ_1 and agent $\nu_J(1)$ is willing to do this because they expect to end up in the matching $\mu_{\nu_J(1)}$.

- The transition from μ_m to μ_{m+1} for some $m \in \{1, \ldots, M-1\}$: Agent m clearly does not want to transition out of the matching μ_m. That means that we need agent $\nu_J(m)$ to break the match $\{m, \nu_J(m)\}$, which the agent is willing to do because they expect to end up in the matching $\mu_{\nu_J(m)}$. Also, we need agents $m+1$ and $\nu_J(m+1)$ to form the match $\{m+1, \nu_J(m+1)\}$. Since, by construction, $\nu_J(m+1) \neq m$, agents $m+1$ and $\nu_J(m+1)$ are willing to transition from μ_m to μ_{m+1} because they believe they will end up being matched to their preferred partner in J. (Note that the agents $\nu_J(m), m+1$, and $\nu_J(m+1)$ are not necessarily distinct. It may be that $\nu_J(m) = m+1$ or $\nu_J(m) = \nu_J(m+1)$.)

- The transition from μ_M to μ_0: Agent $\nu_J(M)$ is effective in this transition and is willing to make it because they expect to end up being matched to their preferred partner in J rather than to agent M.

Select an arbitrary matching $\mu' \neq \mu$ and define the matchings $\mu'^1, \mu'^2, \ldots, \mu'^{K-1}$ as follows:

$$\mu'^k = (\mu^*_{I_1}, \mu^*_{I_2}, \ldots, \mu^*_{I_k}, \mu'_{S_{k+1}})$$

for any $k = 1, \ldots, K-1$. Thus, in the matching μ'^k, all the agents in $I_1 \cup I_2 \cup \ldots \cup I_k$ are in their recursive top-matches, all matches among two agents in $S_{k+1} = N \setminus (I_1 \cup I_2 \cup \ldots \cup I_k)$ are preserved, and agents in S_{k+1} who in μ' were matched to an agent in $I_1 \cup I_2 \cup \ldots \cup I_k$ are single in μ'^k.

Note that if $\mu'(i) = \mu^*_{I_1}(i)$ for each $i \in I_1$, then $\mu'^1 = \mu'$. In this case, the two nodes for μ'^1 and μ' are superimposed in Figure 4. Similarly, for each $k = 2, \ldots, K-1$, if $\mu'^{k-1}(i) = \mu^*_{I_k}(i)$ for each $i \in I_k$, then $\mu'^k = \mu'^{k-1}$ and two nodes for μ'^k and μ'^{k-1} are superimposed in Figure 4. Since K is defined as the smallest number such that either $S_K = \emptyset$ or $I_K = \emptyset$, all agents in I have been matched to their recursive top-matches in μ'^{K-1}. In the matching μ'^{K-1}, the agents in $J = N \setminus I = S_K$ are in their original match in μ' if they were matched to an agent in J, and they are single otherwise.

Note that it is possible that $\mu'^{K-1} = \mu$, or $\mu'^{K-1} = \mu_0$, or $\mu'^{K-1} = \mu_j$ for some $j \in J$. We first consider the general case where μ'^{K-1} is not equal to any of these matchings, which are exactly the matchings in the directed cycle $\mu_0, \mu, \mu_1, \ldots, \mu_M, \mu_0$ that we constructed above. We give each agent in I the path persistent expectations represented by the path $\mu' \to \mu'^1 \to \ldots \to \mu'^{K-1} \to \mu_0 \to \mu$. Also, we extend the expectations of agent $j \in J$ to the matchings $\mu', \mu'^1, \ldots, \mu'^{K-1}$ with the path persistent expectations represented by the path $\mu' \to \mu'^1 \to \ldots \to \mu'^{K-1} \to \mu_0 \to \mu \to \mu_1 \to \ldots \to \mu_j$. Thus, from μ' they expect the path that leads to the cycle and along the cycle they expect the path to terminate at μ_j. (See Figure 4.)

We check that all the transitions that we added in these path persistent expectations are rationalizable:

- The transition from μ' to μ'^1 (only if $\mu'^1 \neq \mu'$): The agents in I_1 for whom $\mu'(i) \neq \mu^*_{I_1}(i)$ are effective in the transition from μ' to μ'^1 and all these agents prefer the matching μ to the matching μ'. Thus, because all agents in I have expectations that they end up in matching μ, the transition from μ' to μ'^1 is rationalizable.

- The transition from μ'^{k-1} to μ'^k for some $k \in \{2, \ldots, K-1\}$ (only if $\mu'^k \neq \mu'^{k-1}$): The agents in I_k for whom $\mu'^{k-1}(i) \neq \mu^*_{I_k}(i)$ are effective in the move from μ'^{k-1} to μ'^k and all these agents prefer the matching μ to the matching μ'^{k-1}. Thus, because all agents in I have expectations that they end up in matching μ, the transition from μ'^{k-1} to μ'^k is rationalizable.

- The transition from μ'^{K-1} to μ_0: Only agents in J are needed in this transition. For every pair of agents $j, k \in J$ that are matched in μ'^{K-1}, it holds that either $j \neq \nu_J(k)$ or $k \neq \nu_J(j)$ (or both). If $j \neq \nu_J(k)$, then agent k is willing to cooperate in the transition from μ'^{K-1} to μ_0 because they expect to end up in the matching μ_k in which they are matched to their preferred partner $\nu_J(k)$ in J. Thus, from each pair of agents that are matched in μ'^{K-1}, we can select an agent who is willing to cooperate in the transition from μ'^{K-1} to μ_0 and the set of all such agents is effective in the move from μ'^{K-1} to μ_0.

- The transition from μ_0 to μ is rationalizable by the expectations of agents in J, as we have demonstrated above.

We have now established that the matching μ i-dominates the matching μ' for each $i \in I$, provided that μ'^{K-1} is not equal to any of the matchings $\mu_0, \mu, \mu_1, \ldots, \mu_M$. We now consider these remaining cases.

If $J = \emptyset$, then $\mu'^{K-1} = \mu$. In this case, we give each player in I the path persistent expectations represented by the path $\mu' \to \mu'^1 \to \ldots \to \mu'^{K-1} (= \mu)$. We have already established that all transitions in these expectations are rationalizable by the expectations of agents in I and thus we can conclude that the matching μ i-dominates the matching μ' for each $i \in I$.

If $\mu_J = (\{j\})_{j \in J}$, then $\mu'^{K-1} = \mu_0$. In this case, the nodes for μ'^{K-1} and μ_0 are superimposed in Figure 4 and the conclusion that the matching μ i-dominates the matching μ' for each $i \in I$ stands.

If $\mu'^{K-1} = \mu_j$ for some $j \in J$, then we give the agents path persistent expectations such that the path $\mu' \to \mu'^1 \to \ldots \to \mu'^{K-1}$ "feeds into the cycle" $\mu_0, \mu, \mu_1, \ldots, \mu_M, \mu_0$ at the matching μ_j rather than the matching μ_0 as follows. We give each agent in I the path persistent expectations represented by the path $\mu' \to \mu'^1 \to \ldots \to \mu'^{K-1} = \mu_j \to \mu_{j+1} \to \ldots \to \mu_M \to \mu_0 \to \mu$. Also, we extend the expectations of agent j to the matchings $\mu', \mu'^1, \ldots, \mu'^{K-1}$ with the path persistent expectations represented by the path $\mu' \to \mu'^1 \to \ldots \to \mu'^{K-1} = \mu_j$. Lastly, we extend the expectations of the agent $k \in J$, $k \neq j$, to the matchings $\mu', \mu'^1, \ldots, \mu'^{K-1}$ with the path persistent expectations represented by the path $\mu' \to \mu'^1 \to \ldots \to \mu'^{K-1} = \mu_j \to \mu_{j+1} \to \ldots \to \mu_k$ (going through μ_0 and μ if $k < j$). Thus, from μ' they expect the path that leads to the cycle and along the cycle they expect the path to terminate at μ_k. We have already established that all transitions in these expectations are rationalizable and thus we can conclude that the matching μ i-dominates the matching μ' for each $i \in I$.

For the arbitrary matching $\mu' \neq \mu$, we have now established that we can find consistent expectations for the agents on the matchings $\mu', \mu'^1, \mu'^2, \ldots, \mu'^{K-1}, \mu_0, \mu, \mu_1, \ldots, \mu_M$ that guarantee that the matching μ i-dominates the matching μ' for each $i \in I$. To establish that $\{\mu\}$ is a singleton farsighted stable set with heterogeneous expectations, we need to demonstrate that we can extend the definitions of these expectations in a consistent manner to *all* matchings in such a way that for each matching $\mu'' \neq \mu$ and for each agent $i \in I$ the agent's expectations from μ'' are a path that terminates at μ. To do this, we can use the same methodology that we described for the matching μ', with the caveat that we do not create more than one node for each matching, but feed a new path into the existing node when we encounter a matching for which we have already created a node. Path persistence is guaranteed because we use the recursive top-matches to create the matchings $\mu', \mu'^1, \mu'^2, \ldots, \mu'^{K-1}, \mu_0, \mu, \mu_1, \ldots, \mu_M$ and the sets $I_1, \ldots, I_{K-1}$ are independent of the matching μ'.

(Necessity) Suppose that μ is a matching such that $\mu(i) \neq \mu^*_I(i)$ for some $i \in I$. We will show that $\{\mu\}$ is not a singleton farsighted stable set with heterogeneous expectations.

Let $k \in \{1, \ldots, K\}$ be the smallest index such that there exists an agent $i \in I_k$ with $\mu(i) \neq \mu^*_{I_k}(i)$. Consider a matching μ' in which all agents in $I_1 \cup \ldots \cup I_k$ are matched according to $(\mu^*_{I_1}, \ldots, \mu^*_{I_k})$. Clearly, $\mu' \neq \mu$. There are no consistent expectations such that at least one agent's expectations at μ' are a path that terminates at μ. This can be seen as follows. No transition out of matching μ' that requires any of the agents in I_1 to be effective is rationalizable because the agents in I_1 do not have any better options. Given that, no transition out of matching μ' that requires any of the agents in I_2 to be effective is rationalizable because the agents in I_2 do not have any better options without the cooperation of the agents in I_1. Continuing in this manner, we establish that no transition out of matching μ' that requires any of the agents in $I_1 \cup \ldots \cup I_k$ to be effective is rationalizable because it is impossible that all the effective agents have expectations such that they are better off at their respective terminal nodes. So, the only agents that can possibly have expectations that support rationalizable transitions out of μ' are the agents not in $I_1 \cup \ldots \cup I_k$. However, these agents are not effective for transitioning to a matching in which the agents in $I_1 \cup \ldots \cup I_k$ are not matched according to $(\mu^*_{I_1}, \ldots, \mu^*_{I_k})$. Therefore, there exist no consistent expectations that support a transition out of $(\mu^*_{I_1}, \ldots, \mu^*_{I_k})$ and thus μ does not i-dominate μ' for any agent i. ■

Proposition 4.8 characterizes the singleton farsighted stable sets with heterogenous expectations when top matches exist. It first asserts that *recursive top matches are necessarily formed in a singleton farsighted stable set.* Hence, if the top-match matching encompasses all agents (i.e., $I = N$), as is the case in Diamantoudi and Xue [3] and Mauleon, Vannetelbosch and Vergote [11], then we get the result that there is a unique singleton farsighted stable set. This unique singleton farsighted stable set has as its unique element the top-match matching μ^*. If not all agents can be matched in recursive top matches (i.e., $J = N \setminus I \neq \emptyset$), then farsighted stable sets with heterogeneous expectations becomes a very permissive solution concept: The agents in I have to be matched according to the top-match matching, but the agents in J can be matched to each other in any way. With μ_J an arbitrary matching of the agents in J, the construction of the indirect dominance paths leading to the matching $\mu = (\mu^*_I, \mu_J)$ makes use of the fact that agents in J have heterogeneous expectations. Each agent in J expects to be matched to their preferred partner in J, and therefore, when the agents in I are all matched according to μ^*_I, the agents in J are all willing to be effective in any transition along a path of matchings that includes the matching μ. The singleton farsighted stable set with heterogeneous expectations $\{\mu\}$ is supported by the expectations of players in I. These players are matched according to the top matching μ^*_I and can entertain any expectations over the matchings formed by agents in J that are rationalizable by the expectations of the agents in J. This enables us to sustain a large range of matchings μ as singleton farsighted stable sets with heterogenous expectations, including matchings in which agents in J are matched to their worst partner in J.

We next consider the situation when no top matches exist in N.

Proposition 4.9 *Suppose that $I = \emptyset$. Then for any matching μ it holds that $\{\mu\}$ is a singleton farsighted stable set with heterogeneous expectations if and only if there exists an agent $i \in N$ such that $\mu(i) \succ_i j$ for all agents j such that $\nu(j) = i$.*

Proof: (Sufficiency) We distinguish between two cases, depending on whether or not in the matching μ any agents are matched to their preferred partner. Let $I(\mu) = \{i \in N \mid \mu(i) = \nu(i)\}$ be the set of such agents.

Case 1. $I(\mu) = \emptyset$**.**

Let agent 1 be an agent for whom $\mu(i) \succ_i j$ for all agents j such that $\nu(j) = i$. Note that it is possible that agent 1 is single in the matching μ if the agent is no other agent's top partner.

We construct consistent expectation functions such that from any matching $\mu' \neq \mu$, agent 1 expects to follow a path ending at μ. Because of path persistence and consistency, we can represent the expectation functions using a graph with transitions between matchings. Figure 5 illustrates the construction of the expectation functions. The matching μ_0 is defined by

$$\mu_0 = (\{j\}_{j\in J})$$

and leaves all agents isolated, i.e., $\mu_0(j) = j$ for all $j \in J$. For each $j \neq 1$, we define the matching

$$\mu_j = (\{j, \nu(j)\}, \{k\}_{k\in J,\, k\neq j, \nu(j)}),$$

in which agent j and their preferred partner $\nu(j)$ are matched to each other and all other agents are single.

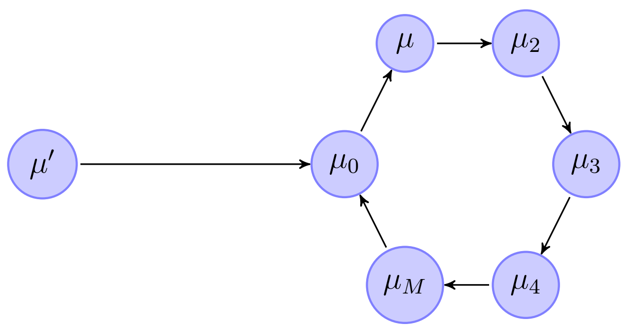

Figure 5: Expectation functions when $I = \emptyset$ and $I(\mu) = \emptyset$

Because $\nu(j) \neq j$ for all $j \in J$, all the matchings μ_j are clearly different from μ_0 and because $\nu(\nu(j)) \neq j$ for all agents in J (by definition of J), all these matchings are also different from one another. Note that it is possible that $\mu_0 = \mu$, but that $\mu_j \neq \mu$ for every $j \neq 1$.

Using Lemma 4.7, we assign the numbers $2, \dots, M$, where $M = |J|$ to the agents in $\{j \in J \mid j \neq 1\}$ in such a way that $\nu(m+1) \neq m$ for all $m = 1, 2, \dots, M-1$, and $\nu(1) \neq M$. This defines a directed cycle $\mu_0, \mu, \mu_2, \mu_3, \dots, \mu_M, \mu_0$.

Select an arbitrary matching $\mu' \neq \mu$. Note that it is possible that $\mu' = \mu_0$, or $\mu' = \mu_j$ for some $j \in \{2, 3, \dots, M\}$. We first consider the general case where μ' is not equal to any of the matchings in the directed cycle $\mu_0, \mu, \mu_2, \mu_3, \dots, \mu_M, \mu_0$.

We give agent 1 the path persistent expectations represented by the paths $\mu' \to \mu_0 \to \mu$ and $\mu_2 \to \mu_3 \to \dots \to \mu_M \to \mu_0 \to \mu$. Thus, from μ' they expect to go to the matching μ_0 in the cycle and from any matching along the cycle they expect a path that follows the cycle and terminates at μ. We give agent $j \in 2, 3, \dots M$ the path persistent expectations represented by the paths $\mu' \to \mu_0 \to \mu \to \mu_2 \to \dots \to \mu_j$ and $\mu_{j+1} \to \mu_{j+2} \to \dots \to \mu_0 \to \mu \to \mu_2 \to \dots \to \mu_j$. Thus, from μ' they expect to go to the matching μ_0 in the cycle and from any matching along the cycle they expect a path that follows the cycle and terminates at μ_j.

We check that all the transitions in these path persistent expectations are rationalizable:

- The transition from μ_0 to μ (only if $\mu_0 \neq \mu$): Every agent $j \in \{2, 3, \ldots, M\}$ is willing to cooperate in the transition from μ_0 to μ because they expect to end up in the matching μ_j in which they are matched to their preferred partner, and they strictly prefer that match to being single. If agent 1 is single in μ, then they are not effective for the transition from μ_0 to μ. If agent 1 is not single in μ, then they are willing to cooperate in the transition from μ_0 to μ because they expect to end up in the matching μ and $\mu(1) \succ_1 1$.

- The transition from μ to μ_2: Note that only agents $j \in \{2, 3, \ldots, M\}$ have expectations to transition from μ to μ_2. Because $I(\mu) = \emptyset$, every player j who is in a matched pair in μ is not matched to their favorite partner. It follows that from every matched pair of agents $\{j, k\}$ in μ, we can select an agent $j \neq 1$ who is willing to cooperate in the transition from μ to μ_2 because they expect to end up in the matching μ_j in which they are matched to their preferred partner $\nu(j)$. For the transition from μ to μ_2, we also need agents 2 and $\nu(2)$ to effective and form the match $\{2, \nu(2)\}$. By construction, $\nu(2) \neq 1$. Because $I(\mu) = \emptyset$, we know that agents 2 and $\nu(2)$ are not matched in μ and also that agent $\nu(2)$ is not matched to their favorite partner $\nu(\nu(2))$ in μ. Thus, both agents 2 and $\nu(2)$ are willing to form the match $\{2, \nu(2)\}$: Agent 2 is willing to do this because they expect to end up in the matching μ_2 and agent $\nu(2)$ is willing to do this because they expect to end up in the matching $\mu_{\nu(2)}$.

- The transition from μ_m to μ_{m+1} for some $m \in \{2, \ldots, M-1\}$: Agent m clearly does not want to transition out of the matching μ_m. That means that we need agent $\nu(m)$ to break the match $\{m, \nu(m)\}$ and we need agents $m+1$ and $\nu(m+1)$ to form the match $\{m+1, \nu(m+1)\}$. By construction, $\nu(m+1) \neq m$, but it may be that $\nu(m) = m+1$ or $\nu(m) = \nu(m+1)$. First, we show that agent $\nu(m)$ is willing to break the match $\{m, \nu(m)\}$ because they expect to end up in a better match: If $\nu(m) \in \{2, 3, \ldots, M\}$, they expect to end up being matched to their preferred partner in the matching $\mu_{\nu(m)}$. If $\nu(m) = 1$, then agent 1 is willing to break the match $\{m, 1\}$ because they expect to end up in the matching μ and $\mu(1) \succ_1 j$ for all agents j such that $\nu(j) = 1$. Next, we show that agents $m+1$ and $\nu(m+1)$ are willing to form the match $\{m+1, \nu(m+1)\}$: If $\nu(m+1) \neq 1$, both agents $m+1$ and $\nu(m+1)$ expect to end up being matched to their preferred partner. If $\nu(m+1) = 1$, then this agent expects to end up in the matching μ. We distinguish between two cases. Case 1: If $\nu(m) = \nu(m+1) = 1$, then we have already shown that agent 1 is willing to be effective in the transition from μ_m to μ_{m+1}. Case 2: If $\nu(m) \neq \nu(m+1) = 1$, then agent 1 is single in μ_m and is willing to be effective in the transition from μ_m to μ_{m+1} because they expect to end up in the matching μ and $\mu(1) \succ_1 m+1 \succ_1 1$.

- The transition from μ_M to μ_0: Agent $\nu(M)$ is effective in this transition. If $\nu(M) \neq 1$, they expect to end up being matched to their preferred partner in the matching $\mu_{\nu(M)}$. If $\nu(M) = 1$, then agent 1 is willing to break the match $\{M, 1\}$ because they expect to end up in the matching μ and $\mu(1) \succ_1 j$ for all agents j such that $\nu(j) = 1$.

- Transition from μ' to μ_0: From every pair of agents that are matched to each other in μ', we need at least one agent j to be effective for this transition. Consider two agents j, k such that $\{j, k\}$ is a match in μ'. Because $I = \emptyset$, either $j \neq \nu(k)$ or $k \neq \nu(j)$ (or both). Thus, from every matched pair of agents $\{j, k\}$ in μ', we can select an agent j such that $\mu'(j) \neq \nu(j)$. If this agent j is different from 1, then they are willing to cooperate in the transition from μ' to μ_0 because they expect to end up in the matching μ_j in which they are matched to their preferred partner $\nu(j)$. If $\mu'(1) = j \neq 1$ and $\nu(j) = 1$, i.e., in μ' agent 1 is matched to an agent for whom they are the preferred partner, then agent 1 is willing to break the match $\{1, j\}$ because they expect to end up in the matching μ and $\mu(1) \succ_1 j$ (since $\nu(j) = 1$).

We have now established that the matching μ 1-dominates the matching μ', provided that μ' is not equal to any of the matchings $\mu_0, \mu_2, \mu_3, \ldots, \mu_M$. If μ' is equal to μ_0 or μ_j for some $j \in \{2, 3, \ldots, M\}$, then one of the nodes in the cycle $\mu_0, \mu, \mu_2, \mu_3, \ldots, \mu_M, \mu_0$ also represents the matching μ' and we simply do not have a separate node for μ' in Figure 5. We have already established that all transitions in the expectations along the cycle are rationalizable and thus we can conclude that the matching μ 1-dominates the matching μ'.

The consistent expectations for the agents on the matchings $\mu', \mu_0, \mu, \mu_2, \mu_3, \ldots, \mu_M$ that we described guarantee that the matching μ 1-dominates each of the matchings μ', μ_0, and $\mu_2, \mu_3, \ldots, \mu_M$. To establish that $\{\mu\}$ is a singleton farsighted stable set with heterogeneous expectations, we need to demonstrate that we can extend the definitions of these expectations in a consistent manner to *all* matchings in such a way that for each matching $\mu'' \neq \mu$ agent 1's expectations from μ'' are a path that terminates at μ. To do this, we simply add a transition from μ'' to μ_0 for each matching μ'' that is not equal to one of the matchings that we already covered.

Case 2. $I(\mu) \neq \emptyset$**.**
We construct slightly different expectations in the following two cases.

Case 2.a. If there exists an agent $i \in I(\mu)$ such that $\nu(\nu(i)) \notin I(\mu)$, then we let 1 be such an agent, and we let 2 be the agent $\nu(1)$. Thus, in the matching μ, 1 is matched to their top partner 2 and, because $I = \emptyset$, it holds that $\nu(2) \neq 1$ and $2 \in J \setminus I(\mu)$. Using Lemma 4.7, we assign the numbers $3, \ldots, M$, where $M = |J \setminus I(\mu)| + 1$ to the agents in

$\{j \in J \mid j \neq 2 \text{ and } \mu(j) \neq \nu(j)\}$ in such a way that $\nu(m+1) \neq m$ for all $m = 2, 3, \ldots, M-1$, and $\nu(1) \neq M$.

Case 2.b. If for each agent $i \in I(\mu)$ it holds that $\nu(\nu(i)) \in I(\mu)$, then select $1 \in I(\mu)$ arbitrarily and let $2 = \nu(1)$. We define a matching $\tilde{\mu}_2$ as follows:

$$\tilde{\mu}_2 = (\{2, \nu(\nu(2))\}, \{k\}_{k \in J,\, k \neq 2, \nu(\nu(2))}),$$

in which agents 2 and $\nu(\nu(2))$ are matched to each other and all other agents are single. Note that $2 = \nu(1) \in J \setminus I(\mu)$ and that $\nu(2) \neq 1$ (because $I = \emptyset$). Also, $\nu(2) = \nu(\nu(1)) \in I(\mu)$, meaning that agents $\nu(2)$ and $\nu(\nu(2))$ are matched to each other in μ. Using that $I = \emptyset$, we thus derive that $\nu(\nu(2)) \neq 2$ and $\nu(\nu(2)) \notin I(\mu)$. Thus, the matching $\tilde{\mu}_2$ exists of exactly one matched pair of agents, and both of those agents are not in $I(\mu)$ and thus not matched to their preferred partner in μ. Also, $\nu(j) \neq j$ for all j implies that $\nu(\nu(2)) \neq \nu(2)$. Moreover, since agent $\nu(2)$ is matched to their preferred partner in μ, by the conditions that define case 2, it holds that $\nu(\nu(\nu(2))) \in I(\mu)$ and thus agent 2, who is not in $I(\mu)$, is not the preferred partner of agent $\nu(\nu(2))$. Hence neither of the two agents 2 and $\nu(\nu(2))$ are matched to their preferred partner in $\tilde{\mu}_2$.

Using Lemma 4.7, we assign the numbers $3, \ldots, M$, where $M = |J \setminus I(\mu)| + 1$ to the agents in $\{j \in J \mid j \neq 2 \text{ and } \mu(j) \neq \nu(j)\}$ in such a way that $\nu(m+1) \neq m$ for all $m = 2, 3, \ldots, M-1$, and $\nu(1) \neq M$.

Cases 2.a and 2.b. Note that in both cases 2.a. and 2.b., it holds that $1 \in I(\mu)$ and thus $I = \emptyset$ implies that $\mu(1) = \nu(1) \notin I(\mu)$, that $\mu(1) \succ_1 1$, and that $\mu(1) \succ_1 j$ for all agents j such that $\nu(j) = 1$. In addition, in both cases it holds that $J \setminus I(\mu) = \{2, 3, \ldots, M\}$.

We construct consistent expectation functions such that from any matching $\mu' \neq \mu$, all agents in $I(\mu)$ expect to follow a path ending at μ. Because of path persistence and consistency, we can represent the expectation functions using a graph with transitions between matchings. Figure 6 illustrates the construction of the expectation functions. The matching μ_0 is defined by

$$\mu_0 = (\{j\}_{j \in J})$$

and leaves all agents isolated, i.e., $\mu_0(j) = j$ for all $j \in J$. For each $j \in J \setminus I(\mu)$, we define the matching

$$\mu_j = (\{j, \nu(j)\}, \{k\}_{k \in J,\, k \neq j, \nu(j)}),$$

in which agent j and their preferred partner $\nu(j)$ are matched to each other and all other agents are single.

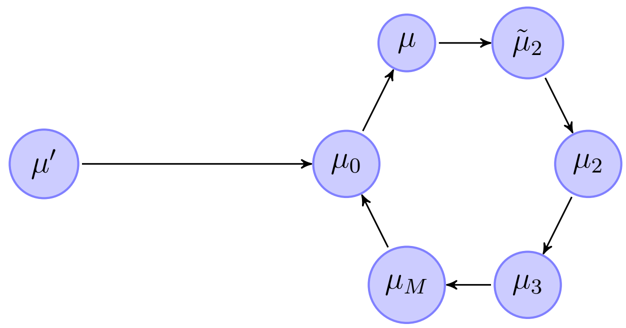

Figure 6: Expectation functions when $I = \emptyset$ and $I(\mu) \neq \emptyset$

Because $\nu(j) \neq j$ for all $j \in J$, all the matchings μ_j are clearly different from μ_0 and because $\nu(\nu(j)) \neq j$ for all agents in J (by definition of J), all these matchings are also different from one another. Note that $1 \in I(\mu)$ implies that agent 1 is matched in μ and thus $\mu_0 \neq \mu$. Also, $\mu_j \neq \mu$ for every $j \in J \setminus I(\mu)$. In case 2.a., we do not need the matching $\tilde{\mu}_2$ and in case 2.b. the matching $\tilde{\mu}_2$ is carefully constructed to be different from each of the matchings μ, μ_0, and μ_j for each $j \in J \setminus I(\mu)$.

In what follows, we cover both cases 2.a. and 2.b. simultaneously and we mention the matching $\tilde{\mu}_2$ with the understanding that we skip $\tilde{\mu}_2$ if the matching μ falls in case 2.a and we include it if μ falls into case 2.b.

Consider the directed cycle $\mu_0, \mu, \tilde{\mu}_2, \mu_2, \mu_3, \ldots, \mu_M, \mu_0$.

Select an arbitrary matching $\mu' \neq \mu$. Note that it is possible that $\mu' = \mu_0$, $\mu' = \tilde{\mu}_2$, or $\mu' = \mu_j$ for some $j \in \{2, 3, \ldots, M\}$. We first consider the general case where μ' is not equal to any of the matchings in the directed cycle $\mu_0, \mu, \tilde{\mu}_2, \mu_2, \mu_3, \ldots, \mu_M, \mu_0$.

We give all the agents in $I(\mu)$ the path persistent expectations represented by the paths $\mu' \to \mu_0 \to \mu$ and $\tilde{\mu}_2 \to \mu_2 \to \mu_3 \to \ldots \to \mu_M \to \mu_0 \to \mu$. Thus, from μ' they expect to go to the matching μ_0 in the cycle and from any matching along the cycle they expect a path that follows the cycle and terminates at μ. We give agent $j \in J \setminus I(\mu) = \{2, 3, \ldots, M\}$ the path persistent expectations represented by the paths $\mu' \to \mu_0 \to \mu \to \tilde{\mu}_2 \to \mu_2 \to \ldots \to \mu_j$ and $\mu_{j+1} \to \mu_{j+2} \to \ldots \to \mu_M \to \mu_0 \to \mu \to \tilde{\mu}_2 \to \mu_2 \to \ldots \to \mu_j$. Thus, from μ' they expect to go to the matching μ_0 in the cycle and from any matching along the cycle they expect a path that follows the cycle and terminates at μ_j.

We check that all the transitions in these path persistent expectations are rationalizable:

- The transition from μ_0 to μ: Every agent $j \in I(\mu)$ is willing to cooperate in the transition from μ_0 to μ because they expect to end up in the matching μ in which they are matched to their preferred partner, and they strictly prefer that match to being single. Every agent $j \in \{2, 3, \ldots, M\}$ is willing to cooperate in the transition from μ_0 to μ because they expect to end up in the matching μ_j in which they are matched to their preferred partner, and they strictly prefer that match to being single.

- The transition from μ to $\tilde{\mu}_2$ (only in case 2.b.): Note that only agents $j \in \{2, 3, \ldots, M\}$ have expectations to transition from μ to $\tilde{\mu}_2$. For the transition from μ to $\tilde{\mu}_2$, we need agents 2 and $\nu(\nu(2))$ to effective, as well as at least one agent $j \in \{2, 3, \ldots, M\}$ from every pair of agents that are matched to each other in μ. First, consider two agents j, k such that $\{j, k\}$ is a match in μ. Because $I = \emptyset$, either $j \neq \nu(k)$ or $k \neq \nu(j)$ (or both). It follows that from every matched pair of agents $\{j, k\}$ in μ, we can select an agent $j \in J \setminus I(\mu)$ such that $\mu(j) \neq \nu(j)$. This agent j is willing to cooperate in the transition from μ to $\tilde{\mu}_2$ because they expect to end up in the matching μ_j in which they are matched to their preferred partner $\nu(j)$. It remains to show that the two agents 2 and $\nu(\nu(2))$ are willing to form the match $\{2, \nu(\nu(2)))\}$. By construction, $2 \notin I(\mu)$ and $\nu(\nu(2)) \notin I(\mu)$, and both these agents are willing to cooperate in the transition from μ to $\tilde{\mu}_2$ because each of them expects to end up being matched to their preferred partner.

- The transition from $\tilde{\mu}_2$ to μ_2 (only in case 2.b.): For this transition, we need agents 2 and $\nu(2)$ to be effective and form the match $\{2, \nu(2)\}$. Agent 2 is matched to $\nu(\nu(2)) \neq \nu(2)$ in $\tilde{\mu}_2$ and is willing to break the match $\{2, \nu(\nu(2))\}$ and form the match $\{2, \nu(2)\}$ because they expect to stay in the matching μ_2. Agent $\nu(2)$ is single in $\tilde{\mu}_2$ and is willing to form the match $\{2, \nu(2)\}$ because they expect to end up in the matching $\mu_{\nu(2)}$.

- The transition from μ to μ_2 (only in case 2.a.): Note that only agents $j \in \{2, 3, \ldots, M\}$ have expectations to transition from μ to μ_2. For the transition from μ to μ_2, we need agents 2 and $\nu(2)$ to effective, as well as at least one agent $j \in \{2, 3, \ldots, M\}$ from every pair of agents that are matched to each other in μ. First, consider two agents j, k such that $\{j, k\}$ is a match in μ. Because $I = \emptyset$, either $j \neq \nu(k)$ or $k \neq \nu(j)$ (or both). It follows that from every matched pair of agents $\{j, k\}$ in μ, we can select an agent $j \in J \setminus I(\mu)$ such that $\mu(j) \neq \nu(j)$. This agent j is willing to cooperate in the transition from μ to μ_2 because they expect to end up in the matching μ_j in which they are matched to their preferred partner $\nu(j)$. It remains to show that the two agents 2 and $\nu(2)$ are willing to form the match $\{2, \nu(2)\}$. By construction, $2 \notin I(\mu)$ and $\nu(2) \notin I(\mu)$, and both these agents are willing to cooperate in the transition from μ to μ_2 because each of them expects to end up being matched to their preferred partner.

- The transition from μ_m to μ_{m+1} for some $m \in \{2, \ldots, M-1\}$: Agent m clearly does not want to transition out of the matching μ_m. That means that we need agent $\nu(m)$ to break the match $\{m, \nu(m)\}$ and we need agents $m+1$ and $\nu(m+1)$ to form the match $\{m+1, \nu(m+1)\}$. By construction, $\nu(m+1) \neq m$, but it may be that $\nu(m) = m+1$ or $\nu(m) = \nu(m+1)$. First, note that agent $\nu(m)$ is willing to break the match $\{m, \nu(m)\}$ because they expect to end up being matched to their preferred partner (in matching $\mu_{\nu(m)}$ if $\nu(m) \in \{2, 3, \ldots, M\}$, and in matching μ if $\nu(m) \in I(\mu)$). Also, agents $m+1$

and $\nu(m+1)$ are willing to form the match $\{m+1, \nu(m+1)\}$: Agent $m+1$ is not matched to their preferred partner in matching μ_m ($\nu(m+1) \neq m$ by construction) and expects to end up in matching μ_{m+1}. Agent $\nu(m+1)$ expects to end up being matched to their preferred partner (in matching $\mu_{\nu(m+1)}$ if $\nu(m+1) \in \{2, 3, \ldots, M\}$, and in matching μ if $\nu(m+1) \in I(\mu)$).

- The transition from μ_M to μ_0: Agent $\nu(M)$ is effective in this transition and is willing to make it because they expect to end up being matched to their preferred partner rather than to agent M.

- Transition from μ' to μ_0. From every pair of agents that are matched to each other in μ', we need at least one agent j to be effective for this transition. Consider two agents j, k such that $\{j, k\}$ is a match in μ'. Because $I = \emptyset$, either $j \neq \nu(k)$ or $k \neq \nu(j)$ (or both). Thus, from every matched pair of agents $\{j, k\}$ in μ', we can select an agent j such that $\mu'(j) \neq \nu(j)$. If this agent j is is not in $I(\mu)$, then they are willing to cooperate in the transition from μ' to μ_0 because they expect to end up in the matching μ_j in which they are matched to their preferred partner $\nu(j)$. If $j \in I(\mu)$, then they are willing to cooperate in the transition from μ' to μ_0 because they expect to end up in the matching μ and they are matched to their preferred partner $\nu(j)$ in that matching.

We have now established that the matching μ i-dominates the matching μ' for each $i \in I(\mu)$, provided that μ' is not equal to any of the matchings $\mu_0, \tilde{\mu}_2, \mu_2, \mu_3, \ldots, \mu_M$. If μ' is equal to μ_0, or $\tilde{\mu}_2$ (only in case 2.b.), or μ_j for some $j \in \{2, 3, \ldots, M\}$, then one of the nodes in the cycle $\mu_0, \mu, \tilde{\mu}_2, \mu_2, \mu_3, \ldots, \mu_M, \mu_0$ also represents the matching μ' and we simply do not have a separate node for μ' in Figure 6. We have already established that all transitions in the expectations along the cycle are rationalizable and thus we can conclude that the matching μ i-dominates the matching μ' for each $i \in I(\mu)$.

The consistent expectations for the agents on the matchings $\mu', \mu_0, \mu, \tilde{\mu}_2, \mu_2, \mu_3, \ldots, \mu_M$ that we described guarantee that the matching μ i-dominates each of the matchings μ', μ_0, $\tilde{\mu}_2$, and $\mu_2, \mu_3, \ldots, \mu_M$, for each $i \in I(\mu)$. To establish that $\{\mu\}$ is a singleton farsighted stable set with heterogeneous expectations, we need to demonstrate that we can extend the definitions of these expectations in a consistent manner to *all* matchings in such a way that for each matching $\mu'' \neq \mu$ agent i's expectations from μ'' are a path that terminates at μ, for each $i \in I(\mu)$. To do this, we simply add a transition from μ'' to μ_0 for each matching μ'' that is not equal to one of the matchings that we already covered.

(Necessity) Suppose that μ is a matching with the property that there does not exist an agent $i \in N$ such that $\mu(i) \succ_i j$ for all agents j such that $\nu(j) = i$. For such a matching μ it holds that for each agent $i \in N$ there exists an agent j such that $i = \nu(j)$ and $j \succeq_i \mu(i)$. We will show that $\{\mu\}$ is not a singleton farsighted stable set with heterogeneous expectations.

The reason is that there do not exist consistent expectations that satisfy the property that there there is an agent whose expectations at μ' are a path that terminates at μ for every matching $\mu' \neq \mu$.

Let $\mathbf{p} = \{p_i\}_{i \in N}$ be expectation functions and $i \in N$ an agent such that $\mu = t(p_i, \mu')$ for each matching μ'. These expectations have to involve transitions that are not rationalizable, and thus they are not consistent. This is seen as follows.

Let j be an agent such that $i = \nu(j)$ and $j \succeq_i \mu(i)$. We prove by contradiction that agents i and j cannot be matched to each other in μ. Suppose that $\mu(j) = i$. Because $I = \emptyset$, we also know that for all agents k for whom j is their favorite partner ($\nu(k) = j$) it is the case that $\nu(j) \neq k$ and thus $\mu(j) = \nu(j) \succ_j k$. This contradicts our assumptions on the matching μ.

Choose a matching μ' in which i and j are matched to each other, i.e., $\mu'(i) = j$ and $\mu'(j) = i$. Because i and j are not matched to each other in μ, somewhere on the path $p_i(\mu')$ there is a transition $\mu'' \to \mu'''$ from a matching μ'' that includes the match $\{i, j\}$ to a matching μ''' that does not include it. Either agent i or agent j needs to be effective in this transition. However, $\mu''(j) = i = \nu(j)$ and agent j cannot have any expectations that rationalize their move out of μ''. That leaves agent i, but $j \succeq_i \mu(i)$ and $j \neq \mu(i)$ imply that that $j \succ_i \mu(i)$, so that $\mu''(i) = j \succ_i \mu(i) = (t(p_i, \mu''))(i)$. Thus, the transition $\mu'' \to \mu'''$ is not rationalizable. ■

Proposition 4.9 shows that, in the absence of top matches, a singleton farsighted stable set with heterogenous expectations must involve at least one player being matched to a partner whom they strictly prefer to any agent whose top choice they are. Note that this condition is satisfied if any agent is matched to their most preferred partner, but that the existence of such an agent is not necessary. The condition is weak in the sense that it only involves the match of one of the agents and the matching of all other agents is unrestricted. The construction of expectations supporting this farsighted stable set is slightly different for the three cases in which either (1) no agent is matched to their most preferred partner, or (2) there is an agent i who is matched to their preferred partner j and the preferred partner of j is not matched to their top choice, or (3) each agent j who is matched to an agent whose top choice they are faces the problem that their top choice $\nu(j)$ is already in the best match possible. The matching μ is supported as a singleton farsighted stable set with heterogenous expectations by the expectations of all the agents who are matched to their preferred partner if such agents exist in μ or, if such agents do not exist, by the expectations of one agent who is matched to a partner whom they strictly prefer to any agent whose top choice they are.

5 Voting

We consider a second application to voting among a set of alternatives. This model is initially due to Moulin and Peleg [12] and has been extended by Abdou and Keiding [1] and Storcken [16]. Agents have strict preferences $\succ_i$ over a set of alternatives A. Any move from an alternative a to an alternative b must be enforced by voting. The underlying power structure is given by a simple game that is characterized by the set of winning coalitions $\mathcal{W}$ which have the power to enforce any transition between alternatives. We assume that the simple game is *monotonic* (i.e., if $S \subset T$ and $S \in \mathcal{W}$, then $T \in \mathcal{W}$) and *proper* (i.e., if $S \in \mathcal{W}$, then $N \setminus S \notin \mathcal{W}$). These assumptions guarantee that the effectivity function is monotonic and that two disjoint coalitions cannot both be winning.

Special cases of monotonic and proper simple games are simple majority games. In a simple majority game, a coalition of agents is winning if and only if it contains a simple majority of all the agents. Hence, denoting by n the number of all agents, the winning coalitions S of agents are those that contain more than half of all the agents (i.e., $|S| \geq \frac{n+1}{2}$ if n is odd, and $|S| \geq \frac{n}{2} + 1$ if n is even). A *Condorcet winner* is defined as an alternative a that satisfies the condition that for each other alternative $b \in A$, the coalition of agents who prefer a to b is a winning coalition in the simple majority game. Thus, a Condorcet alternative would win a two-alternative election against each of the other alternatives using a plurality vote. A Condorcet alternative may not exist, but if such an alternative exists it is clearly unique.

We extend the idea behind Condorcet alternatives to the monotonic and proper simple game characterized by the set of winning coalitions $\mathcal{W}$.

Definition 5.1 *An alternative a is* pairwise winning *if and only if for every alternative $b \in A$, $b \neq a$, there exists a coalition $S \in \mathcal{W}$ such that $a \succ_i b$ for all $i \in S$.*

We study farsighted stability in abstract systems generated by voting situations. Note that in these systems the same coalitions are effective for any transition between alternatives.

We start by showing that under homogeneous expectations, there is a one-to-one correspondence between pairwise winning alternatives and singleton farsighted stable sets.

Proposition 5.2 *In the voting model, $\{a\}$ is a singleton farsighted stable set with homogeneous expectations if and only if a is a pairwise winning alternative.*

Proof: (Sufficiency) Suppose a is a pairwise winning alternative. Then for every alternative $b \neq a$ there exists a coalition $S \in \mathcal{W}$ of players who all prefer alternative a to alternative b. Such a winning coalition S is effective for the transition from b to a and thus $a \operatorname{dom}_S b$. Hence, a directly dominates b for any alternative $b \neq a$ and $\{a\}$ is a farsighted stable set with homogeneous expectations.

(Necessity) Let a be an alternative such that $\{a\}$ is a singleton farsighted stable set with homogeneous expectations. Let $b \in A$, $b \neq a$. Then there exists a sequence of alternatives $b = a_0, a_1, \ldots, a_k = a$ and a sequence of coalitions $S_1, \ldots, S_k$ such that for each $j \in \{1, 2, \ldots, k\}$ it holds that $a_{j-1} \to_{S_j} a_j$ and $a \succ_i a_{j-1}$ for all $i \in S_j$. Because effectivity for transitions between alternatives is given by the simple game, we know that $S_1, \ldots, S_k \in \mathcal{W}$. Note that coalition $S_1 \in \mathcal{W}$ is also effective for the transition from b to a, and that $a \succ_i a_0$ for all $i \in S_1$. We conclude that for every alternative $b \neq a$, there exists a coalition $S \in \mathcal{W}$ such that $a \succ_i b$ for all $i \in S$, meaning that a is a pairwise winning alternative. ■

Proposition 5.2 is driven by the fact that the same coalitions (namely those in $\mathcal{W}$) are effective for any transition between alternatives, and thus direct dominance and farsighted dominance are equivalent when the agents have homogeneous expectations. One implication of the proposition is that a singleton farsighted stable set with homogeneous expectations only exists when the agents' preferences and effectivity for transitions admit a pairwise winning alternative. If a is the pairwise winning alternative, then $\{a\}$ is the unique farsighted stable set with homogeneous expectations, which coincides with the core and the unique stable set of the voting game. However, existence of a pairwise winning alternative is a very strong requirement on preferences and the power structure of the simple game. By contrast, we will establish existence of singleton farsighted stable sets with heterogeneous expectations under very weak conditions.

As in the case of one-to-one matching problems, the characterization of farsighted stable sets with heterogeneous expectations relies on the structure of top alternatives of the agents. $N = \{1, 2, \ldots, n\}$ denotes the set of all agents. Let $\{a_1, \ldots, a_k, \ldots, a_K\}$ be the set of alternatives that are the most preferred alternative of at least one agent and, for each $k \in \{1, 2, \ldots, K\}$, let I_k be the set of agents whose top alternative is a_k. Hence, each agent is a member of exactly one I_k. If all agents agree on the top alternative, then $K = 1$ and $I_1 = N$, the set of all agents. At the other extreme, if each agent has a different top alternative, then $K = n$ (the number of agents) and I_k contains exactly one agent for each $k \in \{1, 2, \ldots, n\}$.

The following proposition characterizes farsighted stable sets with heterogeneous expectations in cases when there are top alternatives that cannot be prevented from being implemented by the agents with different top alternatives, i.e., $N \setminus I_k \notin \mathcal{W}$ for some k.

Proposition 5.3 *Let a_k be an alternative such that $N \setminus I_k \notin \mathcal{W}$.*

1. *If there exists a top alternative a_j, $a_j \neq a_k$, such that $N \setminus I_j \notin \mathcal{W}$, then there is no singleton farsighted stable set with heterogeneous expectations.*

2. *If $N \setminus I_j \in \mathcal{W}$ for all $j \neq k$, then $\{a_k\}$ is the only possible singleton farsighted stable set with heterogeneous expectations.*

3. *If a_k is a pairwise winning alternative and $N \setminus I_j \in \mathcal{W}$ for all $j \neq k$, then $\{a_k\}$ is a farsighted stable set with heterogeneous expectations.*

4. *If $I_k \in \mathcal{W}$, then $\{a_k\}$ is a farsighted stable set with heterogeneous expectations.*

Proof: Because $N \setminus I_k \notin \mathcal{W}$, any transition out of alternative a_k requires the vote of at least one member of I_k. However, a_k is the top alternative of all agents in I_k and thus, there is no rationalizable transition out of alternative a_k. Therefore, any farsighted stable set with heterogeneous expectations has to include alternative a_k.

Part 1. Let $j \neq k$ be such that $N \setminus I_j \notin \mathcal{W}$. Then any transition out of alternative a_j requires the vote of at least one member of I_j and there is no rationalizable transition out of alternative a_j. Therefore, any farsighted stable set with heterogeneous expectations has to include alternative a_j in addition to alternative a_k.

Part 2. Suppose $N \setminus I_j \in \mathcal{W}$ for all $j \neq k$. Because any farsighted stable set with heterogeneous expectations has to include alternative a_k, the only possible singleton farsighted stable set with heterogeneous expectations is $\{a_k\}$. The set $\{a_k\}$ is a singleton farsighted stable set if and only if there exist rationalizable expectations by the agents in I_k from any $a \neq a_k$ to a_k. There is no simple condition for the existence of such rationalizable expectations, but the following two cases exhibit two separate sufficient conditions for existence.

Part 3. Suppose a_k is a pairwise winning alternative and $N \setminus I_j \in \mathcal{W}$ for all $j \neq k$. Then for each alternative $a \neq a_k$, there is a coalition $S_a \in \mathcal{W}$ of agents who each prefer a_k to a. Monotonicity of the simple game implies that $S_a \cup I_k \in \mathcal{W}$. Thus, the expectations $a \to_{S_a \cup I_k} a_k$ are rationalizable and alternative a_k dominates alternative a through expectations of the agents in I_k. This establishes that $\{a_k\}$ is a singleton farsighted stable set with heterogeneous expectations.

Part 4. Note that $I_k \in \mathcal{W}$ does not necessarily follow from $N \setminus I_k \notin \mathcal{W}$, but that $N \setminus I_k \notin \mathcal{W}$ is implied by $I_k \in \mathcal{W}$ by properness of the simple game. Because for any $j \neq k$ it holds that $I_k \subset N \setminus I_j$, monotonicity of the simple game further implies that $N \setminus I_j \in \mathcal{W}$. Thus, by part 2, $\{a_k\}$ is the only possible singleton farsighted stable set with heterogeneous expectations. Because for each alternative $a \neq a_k$ the coalition I_k is effective for the transition from a to a_k and because a_k is the top alternative of all agents in I_k, alternative a_k dominates alternative a through expectations of the agents in I_k. This establishes that $\{a_k\}$ is a singleton farsighted stable set with heterogeneous expectations. ■

Proposition 5.4 characterizes farsighted stable sets with heterogeneous expectations in cases when each top alternative can be prevented from being implemented by the agents with different top alternatives, i.e., $N \setminus I_k \in \mathcal{W}$ for each k. For each agent $i \in N$, we define $\mathcal{K}_i = \{k \mid N \setminus (I_k \cup \{i\}) \notin \mathcal{W}\}$ and $\mathcal{A}_i = \{a \in A \mid a \succ_i a_k \text{ for each } k \in \mathcal{K}_i\}$. Thus, $\mathcal{K}_i$ is the set of indices of top alternatives for which agent i is needed in order to prevent them from being implemented, and $\mathcal{A}_i$ is the set of alternatives that agent i prefers to those. Note that $N \setminus I_k \in \mathcal{W}$ for each k implies that $i \notin I_k$ for each $k \in \mathcal{K}_i$. It follows that $\mathcal{A}_i$ contains i's top alternative.

Proposition 5.4 *Suppose that $N \setminus I_k \in \mathcal{W}$ for all $k = 1, 2, \ldots, K$.*

1. *If there exists an agent $i \in N$ such that $\mathcal{K}_i = \emptyset$, then $\{a\}$ is a singleton farsighted stable set with heterogeneous expectations for any alternative $a \in A$.*

2. *If $\mathcal{K}_i \neq \emptyset$ for each $i \in N$, then $\{a\}$ is a singleton farsighted stable set with heterogeneous expectations if and only if $a \in \bigcup_i \mathcal{A}_i \cup \{a_1, ..., a_K\}$.*

Proof: Because $N \setminus I_k \in \mathcal{W}$ for all $k = 1, 2, \ldots, K$, it follows by monotonicity of the simple game that $N \setminus \{i\} \in \mathcal{W}$ for every $i \in N$.

(Sufficiency) We distinguish between two cases.

Part 1. Let $i \in N$ be an agent such that $\mathcal{K}_i = \emptyset$ and let $a \in A$ be an arbitrary alternative. Number the elements of A as $a^1, a^2, \ldots, a^{|A|}$ and give each agent the path persistent expectations represented by the cycle $a^1 \to a^2 \to \ldots \to a^{|A|} \to a^1$, with $t(p_i, a') = a$ for all alternatives a' and $t(p_j, a') = a_k$ for all alternatives a' and each agent $j \in I_k \setminus \{i\}$. Thus, from each alternative a', agent i expects to end up at alternative a and all other agents expect to end up at their most preferred alternative.

We check that all the transitions in these path persistent expectations are rationalizable: For any alternative a' that is not any agent's top alternative ($a' \neq a_k$ for some k), the agents in $N \setminus \{i\}$ are effective to transition to any other alternative and these agents are all willing to make this transition because they each expect to end up in their most preferred alternative. If $a' = a_k$ for some k, then all agents in $N \setminus (I_k \cup \{i\})$ are willing to transition out of alternative a' because they each expect to end up in their most preferred alternative. Moreover, $\mathcal{K}_i = \emptyset$ so that $N \setminus (I_k \cup \{i\}) \in \mathcal{W}$ and thus $N \setminus (I_k \cup \{i\})$ is effective for the transition out of a' to any other alternative.

We have now established that alternative a i-dominates each other alternative a' and thus $\{a\}$ is a singleton farsighted stable set with heterogeneous expectations.

Part 2. Suppose that $\mathcal{K}_i \neq \emptyset$ for each $i \in N$. Let $a \in \bigcup_i \mathcal{A}_i \cup \{a_1, ..., a_K\}$. We distinguish between two cases. **Case 2.a.** If $a = a_l$ for some l, then choose an agent $i \in I_l$, for whom a is the most preferred alternative. **Case 2.b.** If $a \in \mathcal{A}_i$ for some i, then choose an agent i such that alternative a is contained in $\mathcal{A}_i$.

Cases 2.a and 2.b. Number the elements of A as $a^1, a^2, \ldots, a^{|A|}$ and give each agent the path persistent expectations represented by the cycle $a^1 \to a^2 \to \ldots \to a^{|A|} \to a^1$, with $t(p_i, a') = a$ for all alternatives a' and $t(p_j, a') = a_k$ for all alternatives a' and each agent $j \in I_k \setminus \{i\}$. Thus, from each alternative a', agent i expects to end up at alternative a and all other agents expect to end up at their most preferred alternative.

We check that all the transitions in these path persistent expectations are rationalizable: Let $a' \neq a$. If alternative a' is not any agent's top alternative ($a' \neq a_k$ for some k), the agents in $N \setminus \{i\}$ are effective to transition to any other alternative and these agents are all willing to make this transition because they each expect to end up in their most preferred alternative. If $a' = a_k$ for some k, then we distinguish between cases 2.a and 2.b again.

Case 2.a. If $a = a_l$, then $a' \neq a$ guarantees that $k \neq l$. Then also $i \in I_l$ and thus $i \notin I_k$. All agents in $N \setminus I_k$ are willing to transition out of alternative a' because they each expect to end up in their most preferred alternative. Because $N \setminus I_k \in \mathcal{W}$, the agents in $N \setminus I_k$ are also effective for the transition out of alternative a'.

Case 2.b. If $a \in \mathcal{A}_i$, all agents in $N \setminus (I_k \cup \{i\})$ are willing to transition out of alternative a' because they each expect to end up in their most preferred alternative. If $k \notin \mathcal{K}_i$, then $N \setminus (I_k \cup \{i\}) \in \mathcal{W}$ and $N \setminus (I_k \cup \{i\})$ is effective for the transition out of a_k. If $k \in \mathcal{K}_i$, then $N \setminus (I_k \cup \{i\}) \notin \mathcal{W}$ and agent i is needed to transition out of alternative a_k. Agent i is willing to transition out of alternative a_k because they expect to end up in alternative $a \in \mathcal{A}_i$ and $k \in \mathcal{K}_i$ implies that $a \succ_i a_k$.

We have now established that alternative a i-dominates each other alternative a' and thus $\{a\}$ is a singleton farsighted stable set with heterogeneous expectations.

(Necessity) Suppose that $\mathcal{K}_i \neq \emptyset$ for each $i \in N$ and $a \notin \bigcup_i \mathcal{A}_i \cup \{a_1, ..., a_K\}$. We will show that $\{a\}$ is not a singleton farsighted stable set with heterogeneous expectations. The reason is that there do not exist consistent expectations that satisfy the property that there there is an agent whose expectations at a' are a path that terminates at a for every alternative $a' \neq a$.

Let $\mathbf{p} = \{p_i\}_{i \in N}$ be expectation functions and $i \in N$ an agent such that $a = t(p_i, a')$ for each alternative a'. These expectations have to involve transitions that are not rationalizable, and thus they are not consistent. This is seen as follows. Because $a \notin \mathcal{A}_i$, we can choose a $k \in \mathcal{K}_i$ such that $a_k \succeq_i a$. Because $a \neq a_l$ for any l, it holds that $a_k \succ_i a$. Thus, the transition from a_k to some other alternative cannot involve agent i, who expects to end up in alternative a. That leaves the agents in $N \setminus (I_k \cup \{i\})$ to accomplish the transition out of alternative a_k, but $N \setminus (I_k \cup \{i\}) \notin \mathcal{W}$ because $k \in \mathcal{K}_i$. Hene, the agents in $N \setminus (I_k \cup \{i\})$ are not effective for the transition out of alternative a_k to any other alternative. We conclude that there is no rationalizable transition from alternative a_k to any other alternative and therefore this alternative is not i-dominated by a. ■

Propositions 5.3 and 5.4 show that allowing for heterogeneous expectations greatly enlarges the set of alternatives that can be supported as singleton farsighted stable sets. If one agent is not needed in any of the transitions out of the preferred alternatives of the other agents, then any alternative can be supported as a singleton farsighted stable set.

Propositions 5.3 and 5.4 give a very crisp result when the underlying simple game is a symmetric majority game with quota t. If there exists one set I_k such that $|I_k| \geq n - t + 1$ then $N \setminus I_j \notin \mathcal{W}$ and the only candidate for a singleton farsighted stable set is the favorite alternative of the agents in I_k, a_k. If $|I_j| < n - t + 1$ but there exist at least one set such that $|I_j| = n - t$, then for all agents i who do not belong to any of these sets, $\mathcal{J}(i) = \{j||I_j| = n-t\}$. Then, in addition to the favorite alternatives of the agents, any alternative a such that there exists i for whom $a \succ_i a_j \forall j \in calJ(i)$ can be supported as a singleton farsighted stable set. Finally if $|I_j| < n - t$ for all j, then any alternative can be supported as a farsighted stable set.

We thus can use Propositions 5.3 and 5.4 to characterize singleton farsighted stable sets in the two extreme symmetric majority voting games. In the unanimity game, $t = n$ and hence the only situation where a singleton stable set exists is when all agents agree on the optimal alternative. In the simple majority game, $t = \frac{n}{2} + 1$, if no alternative is preferred by a group of more than $\frac{n}{2} - 1$ agents, all alternatives can be supported. If one alternative a_j is preferred by a group of $\frac{n}{2} - 1$ agents, then all preferred alternatives as well as any alternative a that any agent outside I_j prefers to alternative a_j can be supported. Finally if there is a group of at least $\frac{n}{2}$ agents who share the same preferred alternative a_j, then a_j is the only candidate for a farsighted stable set.

6 Conclusion

This paper analyzes farsighted stable sets when agents have heterogeneous expectations over the dominance paths. We consider expectations functions satisfying two properties of path-persistence and consistency. We show that farsighted stable sets always exist. Any singleton farsighted stable set with common expectations is a farsighted stable set with heterogeneous expectations, but nonsingleton farsighted stable sets in the classical Harsanyi-Chwe sense are not necessarily stable sets with heterogeneous expectations..

We observe that, in the one-to-one matching and voting models, any agent who expects a dominance path to terminate at their favorite state is willing to move along the path. This observation then implies that when all other agents expect to reach their favorite state, one agent can have an expectation function supporting a very large number of states as terminal nodes of the dominance path. Given that we only require one agent to hold expectations supporting a state as a farsighted stable set, the relaxation of the hypothesis of common expectations greatly expands the set of states which can be supported as singleton stable sets. In fact, when a top match exists in the one-to-one matching model, any matching of the other agents (including matchings which are Pareto dominates or not individually rational) can be supported. In the voting model, when agents have different favorite alternatives and there is no veto player, any alternative can be supported.

We interpret these results showing that "anything goes" when agents hold heterogeneous expectations about the dominance path, as an indication that most results in farsighted stability rest on the assumption that agents hold common expectations and perfectly coordinate their moves along dominance paths. In applications, the usefulness of an analysis based on agents' farsighted behavior thus depends on whether agents are able to coordinate their expectations about the sequences of moves of other agents.

References

[1] J Abdou and H Keiding. Effectivity functions in social choice, volume 8 of theory and decision library. series c: Game theory. *Mathematical Programming and Operations Research. Kluwer Academic Publishers Group, Dordrecht*, 1991.

[2] Michael Suk-Young Chwe. Farsighted coalitional stability. *Journal of Economic theory*, 63(2):299–325, 1994.

[3] Effrosyni Diamantoudi and Licun Xue. Farsighted stability in hedonic games. *Social Choice and Welfare*, 21(1):39–61, 2003.

[4] Bhaskar Dutta and Rajiv Vohra. Rational expectations and farsighted stability. 2015.

[5] Lars Ehlers. von neumann–morgenstern stable sets in matching problems. *Journal of Economic Theory*, 134(1):537–547, 2007.

[6] John C Harsanyi. An equilibrium-point interpretation of stable sets and a proposed alternative definition. *Management Science*, 20(11):1472–1495, 1974.

[7] P Herings, Ana Mauleon, and Vincent J Vannetelbosch. Stability of networks under level-k farsightedness. *Ana and Vannetelbosch, Vincent J., Stability of Networks Under Level-K Farsightedness (July 10, 2014)*, 2014.

[8] James S Jordan. Pillage and property. *Journal of Economic Theory*, 131(1):26–44, 2006.

[9] Hideo Konishi and Debraj Ray. Coalition formation as a dynamic process. *Journal of Economic Theory*, 110(1):1–41, 2003.

[10] Ana Mauleon and Vincent Vannetelbosch. Farsightedness and cautiousness in coalition formation games with positive spillovers. *Theory and Decision*, 56(3):291–324, 2004.

[11] Ana Mauleon, Vincent J Vannetelbosch, and Wouter Vergote. von neumann–morgenstern farsightedly stable sets in two-sided matching. *Theoretical Economics*, 6(3):499–521, 2011.

[12] Herve Moulin and Bezalel Peleg. Cores of effectivity functions and implementation theory. *Journal of Mathematical Economics*, 10(1):115–145, 1982.

[13] J von Neumann, Oskar Morgenstern, et al. *Theory of games and economic behavior*, volume 60. Princeton university press Princeton, 1944.

[14] Debraj Ray and Rajiv Vohra. The farsighted stable set. *Econometrica*, 83(3):977–1011, 2015.

[15] Alvin E Roth and Marilda A Oliveira Sotomayor. *Two-sided matching*. Cambridge University Press, Cambridge, 1990.

[16] Ton Storcken. Effectivity functions and simple games. *International Journal of Game Theory*, 26(2):235–248, 1997.

[17] Licun Xue. Coalitional stability under perfect foresight. *Economic Theory*, 11(3):603–627, 1998.