

# MAXIMALITY IN THE FARSIGHTED STABLE SET

Rajiv Vohra  Debraj Ray<sup>†</sup>

March 2019

*Econometrica* **87**, 1763–1779 (2019)

**Abstract.** Harsanyi (1974) and Ray and Vohra (2015) extended the stable set of von Neumann and Morgenstern to impose farsighted credibility on coalitional deviations. But the resulting farsighted stable set suffers from a conceptual drawback: while coalitional moves improve on existing outcomes, coalitions might do *even better* by moving elsewhere. Or other coalitions might intervene to impose their favored moves. We show that every farsighted stable set satisfying some reasonable and easily verifiable properties is unaffected by the imposition of these stringent maximality constraints. The properties we describe are satisfied by many, but not all farsighted stable sets.

KEYWORDS: stable sets, farsightedness, maximality, history dependent expectations.

JEL CLASSIFICATION: C71, D72, D74

## 1. INTRODUCTION


The core identifies payoff profiles that no group, or coalition, can dominate with an allocation that is feasible for the coalition in question. But this classical solution does not ask if the new allocation itself is credibly threatened or “blocked” by *other* coalitions. The problem is that the definition of credibility is often circular — an allocation is credible if it is not challenged by a credible allocation.<sup>1</sup> The vNM stable set (von Neumann and Morgenstern, 1944) cuts through that circularity. Say that a payoff profile is dominated by another profile if some coalition prefers the latter profile and can unilaterally implement the piece of the new profile that pertains to it. A set of feasible payoff profiles  $Z$  is *stable* if it satisfies two properties:

*Internal Stability.* If  $u \in Z$ , it is not dominated by  $u' \in Z$ .

*External Stability.* If  $u \notin Z$ , then there exists  $u' \in Z$  which dominates  $u$ .

Internal and external stability work in tandem to get around the circularity, allowing us to view  $Z$  as a norm or a “standard of behavior” (Greenberg, 1990). No two allocations

---

<sup>†</sup>Ray: New York University and University of Warwick, debraj.ray@nyu.edu; Vohra: Brown University, rajiv\_vohra@brown.edu. Ray acknowledges funding from the National Science Foundation under grant SES-1629370. The names on this paper are in certified random order, as proposed in Ray  Robson (2018). Unfortunately, the Econometric Society does not endorse the use of the random order symbol, thereby forcing us to return to alphabetical order in the published version. We are grateful to four anonymous referees and a Co-Editor for their comments.

<sup>1</sup>Concepts such as the bargaining set (Aumann and Maschler 1964), which only try to build in an additional “round” of domination, are not up to the task.

in the standard threaten each other, and jointly, the standard allocations dominate all non-standard allocations. The relevant solution concept is therefore not *a* payoff profile, but a *set* of payoff profiles that work in unison. It is a beautiful definition.

Yet, temporarily setting beauty aside, there are at least three problems with the concept:

1. *Harsanyi critique*. Suppose that  $u'$  dominates  $u \in Z$ , and that  $u'$  is in turn dominated by  $u'' \in Z$ , as required by vNM stability. Then it is true that  $u'$  isn't "credible," but so what? What if the coalition that proposes  $u'$  only does so to induce  $u''$  in the first place, where it *is* better off? Harsanyi (1974) went on to propose a "farsighted version" of vNM stability, one that permits a coalition to anticipate a chain reaction of payoff profiles, and asking for a payoff improvement at the terminal node of this chain.<sup>2</sup>

2. *Ray-Vohra critique*. Ray and Vohra (2015) highlight a seemingly innocuous device adopted by von Neumann and Morgenstern. Dominance is defined over entire profiles of payoffs. As described above, profile  $u'$  dominates  $u$  when some coalition is better off under  $u'$  and can implement *its piece* of  $u'$  unilaterally. But what about the rest of  $u'$ , which involves allocations of payoffs to others who have nothing to do with the coalition in question? Who allocates these payoffs, and what incentive do they have to comply with the stipulated amounts? To this, von Neumann and Morgenstern would answer that it does not matter: payoffs to outsiders are irrelevant, and only a device for tracking all profiles in a common space. However, once the solution is modified along the lines of Harsanyi, the critique *does* matter: the payoffs accruing to others will fundamentally affect the chain reaction that follows. Their determination cannot be finessed.

3. *Maximality problem*. Domination requires every coalition participating in the chain reaction of proposals and counter-proposals to be better off (relative to their starting points) once the process terminates. But it does not require coalitions to choose their *best* moves (Ray and Vohra 2014, Dutta and Vohra 2017), and it rules out possibly unwelcome interventions by other coalitions. This is a concern everywhere along the entire farsighted blocking chain. That chain is supported by the anticipation that later coalitions participating in the chain will also be "better off" doing so. But now "better off" isn't good enough: what if they gain *even* more by doing something else, and that something else isn't good for the original deviator? Or what if a different coalition intervenes? Faced with such potential complexities, the entire chain of proposals becomes suspect.

This third issue, which we call the maximality problem, forms the subject of our paper.

To fix ideas, consider the following example (Example 5.8, Ray and Vohra 2014). There are two players, 1 and 2, and four states,  $a, b, c, d$ . The payoff profiles by state are

---

<sup>2</sup>There is now a sizable literature the studies farsighted stability in coalitional games. This includes Aumann and Myerson (1988), Chwe (1994), Xue (1998), Diamantoudi and Xue (2003), Herings, Mauleon and Vannetelbosch (2004, 2017), Jordan (2006), Ray (2007), Mauleon, Vannetelbosch and Vergote (2011), Kimya (2015), Ray and Vohra (2015), Bloch and van den Nouweland (2017), Dutta and Vohra (2017) and Dutta and Vartiainen (2018).

$u(a) = (1, 1)$ ,  $u(b) = (0, 0)$ ,  $u(c) = (10, 10)$  and  $u(d) = (0, 20)$ . Suppose that state  $a$  can only be changed by player 1, and that she can only move to  $b$ . From  $b$ , only player 2 can move, and she can move either to  $c$  or  $d$ , both of which are terminal states (no further move is possible from  $c$  or  $d$ ). We claim that the unique farsighted stable set is  $\{c, d\}$ . Certainly, both  $c$  and  $d$  must be in every farsighted stable set. But then,  $a$  and  $b$  are not in any farsighted stable set: the state  $b$  is trivially eliminated, while  $a$  is dominated by a move by player 1 to  $b$  followed by a move by player 2 to  $c$ ; player 1 gains by replacing  $a$  with  $c$  and player 2 gains by replacing  $b$  with  $c$ . But the elimination of  $a$  violates maximality: at  $b$ , player 2's optimal move is to  $d$  rather than to  $c$ . If player 1 were to forecast that, it wouldn't be in her interest to move, making  $a$  a "stable" state.

This example suggests that something like subgame perfection needs to be grafted on to farsighted stability. But cooperative game theory attempts to model free-form negotiations. There is no protocol that sets the "rules of the game," describing who moves (even stochastically) at each node. Noncooperative game theory imposes such protocols,<sup>3</sup> but the gain in precision is in part illusory, for the answers can be notoriously sensitive to the choice of the extensive form. In contrast, the theory of blocking is more open-ended: *any* coalition can move at any stage. In particular, the problem of maximality is not just restricted to the coalition that actually moves, but also applies to other coalitions that could *potentially* move. So, while maximality is related to sequential rationality or subgame perfection, it goes beyond that. In fact, different definitions of maximality are possible depending on which coalitions are "allowed" to move at any state.

The weakest of these, referred to as (just) maximality by Dutta and Vohra (2017), requires only that the moving coalition lacks a better alternative to its stipulated move. A stricter version, strong maximality, rules out deviations by any coalition that intersects the coalition stipulated to move. But in this paper, we take on board the strictest variant: one that asks for immunity to all deviations, not just by the coalition that moves "in equilibrium," or by all those that intersect it, but by *any* coalition. To distinguish this concept from weaker notions of maximality we refer to it as *absolute maximality*.

These variants are not particularly germane to the example above, because the unique farsighted stable set identified there does not satisfy any reasonable notion of maximality. But in a negotiation setting, states are not connected by a highly restricted, tree-like structure describing possible moves. States are combinations of coalition structures and proposed payoff allocations. While it is true that not all coalitions are capable of precipitating one state from another, it is possible to travel from any state to any other. Our main result shows that in the context of negotiations, the example above is an outlier: *every farsighted stable set satisfying reasonable and easily verifiable properties is unaffected by the imposition of absolute maximality* (Theorem 1). These properties are described as A and B in Section 3.1. The theorem is useful because the identification of farsighted

---

<sup>3</sup>For noncooperative approaches to coalition formation, see Chatterjee, Dutta, Ray and Sengupta (1993), Bloch (1996), Okada (1996), Ray and Vohra (1999), and the survey in Ray and Vohra (2014).

stable sets, or even stable sets, is not always an easy task. Having to check if they satisfy maximality adds an additional layer of complexity. It would be extremely desirable if such a check could be sidestepped, and Properties A and B allow us to do just that.

There are several cases of special interest in which it is easy to verify that Properties A and B are satisfied. For instance, any farsighted stable set with a unique payoff profile satisfies both properties. Theorem 2 of Ray and Vohra (2015) shows that such sets always exist in games that possess certain core payoffs termed *separable* allocations (defined in Section 3.2 below). We present a new application of separability: under certain conditions, every competitive equilibrium of an exchange economy is separable and is therefore a single-payoff, absolutely maximal farsighted stable set.

*Simple games* are widely employed in applications to political economy. As Shapley (1962) observed, “a surprising number of the multiperson games found in practice are simple.” For such games, Property B is automatically satisfied by every farsighted stable set. Moreover, we show that under mild restrictions, such games possess farsighted stable sets that also satisfy Property A. Consequently an absolutely maximal farsighted stable set always exists in such games.

As already noted, in an abstract setting a farsighted stable set may not satisfy even the weakest form of maximality. Dutta and Vohra (2017, Example 5 and footnote 12) shows that a farsighted stable set may satisfy maximality but not strong maximality, or satisfy strong maximality but not absolute maximality.<sup>4</sup> But coalitional games have more structure and as our positive results show, that allows us to establish *absolute* maximality in a variety of cases.<sup>5</sup> That said, Properties A and B need not always be satisfied, even in coalitional games. Section 3.4 provides three examples. In Example 1, Property B is satisfied for a farsighted stable set but not A; in Example 2, Property A is satisfied for the set but not B. In either example, the farsighted stable set fails to be absolutely maximal, demonstrating that Theorem 1 is tight. At the same time, while sufficient for absolute maximality, Example 3 shows that these properties are not necessary.

## 2. MAXIMAL FARSIGHTED STABILITY

**2.1. Coalitional Games.** A coalitional game is described by a finite set  $N$  of players and a “characteristic function”  $V$  that assigns to each nonempty subset  $S$  (a “coalition”) a nonempty, closed set of feasible payoff vectors  $V(S)$ . Normalize so that singletons obtain zero and assume that  $V(S) \cap \mathbb{R}_+^S$  is bounded.

---

<sup>4</sup>Although they are concerned with history independent processes, these examples also apply to history dependent processes.

<sup>5</sup>Our positive conclusions are even sharper if we require only maximality, rather than absolute maximality. In that case, as shown in an earlier version of this paper (Vohra (©) Ray 2018), Property B can be dispensed with completely. Moreover, for simple games, even Property A can be dropped.

**2.2. States and Effectivity.** A *state* is a partition or coalition structure of  $N$ , along with a payoff profile  $u$  feasible for that structure. A typical state  $x$  is therefore a pair  $(\pi, u)$  (or  $\{\pi(x), u(x)\}$  when we need to be explicit), where  $u_S \in V(S)$  for each  $S \in \pi$ . Let  $X$  be the set of all states. An *effectivity correspondence*  $E(x, y)$  specifies for each pair of states  $x$  and  $y$  the collection of coalitions that have the power to change  $x$  to  $y$ . Ray and Vohra (2015) argue that effectivity correspondences must satisfy natural restrictions for the relevant solution concepts to make sense. Specifically, we assume throughout:

(E.1) If  $S \in E(x, y)$ ,  $T \in \pi(x)$  and  $T \cap S = \emptyset$ , then  $T \in \pi(y)$  and  $u_T(x) = u_T(y)$ .

(E.2) For every state  $x$ , coalition  $S$ , partition  $\mu$  of  $S$  and payoff  $v \in \mathbb{R}^{|S|}$  with  $v_W \in V(W)$  for each  $W \in \mu$ , there is  $y \in X$  such that  $S \in E(x, y)$ ,  $\mu \subseteq \pi(y)$  and  $u_S(y) = v$ .

Condition E.1 grants some sovereignty to untouched coalitions: by forming,  $S$  cannot directly influence the membership or payoffs of coalitions in the original structure that are entirely unrelated to  $S$ . Condition E.2 grants some sovereignty to moving coalitions: if  $S$  wants to move from  $x$ , it can do so by reorganizing itself (breaking up into smaller pieces if it so wishes, captured by the sub-structure  $\mu$ ), provided that the resulting payoff to it,  $v$ , is feasible ( $v_W \in V(W)$  for every  $W \in \mu$ ). What happens “elsewhere,” however, is not under its control (see, for instance, the sovereignty restriction E.1), which is why E.2 only asserts the existence of *some* state  $y$  satisfying the sovereignty conditions.

**2.3. Farsighted Stability.** A *chain* is a finite collection of states  $\{y^0, y^1, \dots, y^m\}$  and coalitions  $\{S^1, \dots, S^m\}$ , such that for every  $k \geq 1$ , we have  $y^{k-1} \neq y^k$ , and  $S^k$  is effective in moving the state from  $y^{k-1}$  to  $y^k$ :  $S^k \in E(y^{k-1}, y^k)$ . A state  $y$  *farsightedly dominates*  $x$  if there is a chain with  $y^0 = x$  and  $y^m = y$  such that for all  $k = 1, \dots, m$ ,  $u(y)_{S^k} \gg u(y^{k-1})_{S^k}$ . This associated chain will be called a *blocking chain*.

A set of states  $F \subseteq X$  is a *farsighted stable set* if it satisfies two conditions:

*Internal Farsighted Stability.* No state in  $F$  farsightedly dominates another state in  $F$ ;

*External Farsighted Stability.* A state not in  $F$  is farsightedly dominated by a state in  $F$ .

Observe that farsightedness fails to impose any optimization on coalitional moves, barring the requirement that coalitions must be eventually better off participating in the chain rather than not participating at all. Maximality addresses this issue.

**2.4. Absolutely Maximal Farsighted Stable Sets.** To examine maximality for a farsighted stable set, we “embed” that set into an ambient history-dependent *negotiation process*. A *history*  $h$  is a finite sequence of states, along with the coalitions that generate any state transitions. If there is no change of state, the empty coalition is recorded. An *initial history* is just a single state.<sup>6</sup> Let  $x(h)$  be the last state in history  $h$ . A *negotiation*

<sup>6</sup>For instance, players might all begin the negotiation process as standalone singletons, or it may be that some going arrangement or state is already in place.

*process* is a map  $\sigma$  from histories to the new outcome. For each  $h$ ,  $\sigma(h) = \{y(h), S(h)\}$ , where  $y(h)$  is the state that follows  $x(h)$  and  $S(h) \in E(x(h), y(h))$  is the coalition implementing the change. (If  $x(h) = y(h)$ , then  $S(h)$  is empty; i.e., “nothing happens.”) In this way, given any history  $h$ ,  $\sigma$  induces a *continuation chain*.

A state  $x$  is *absorbing* under  $\sigma$  if at any  $h$  with  $x(h) = x$ ,  $y(h) = x(h) = x$ . That is, once at  $x$  the continuation chain displays  $x$  forever. Say that  $\sigma$  is *absorbing* if after every history, its continuation chain ends in an absorbing state.<sup>7</sup> For every absorbing process  $\sigma$  and history  $h$ , let  $x^\sigma(h)$  denote the absorbing state reached from  $h$ . An absorbing process  $\sigma$  is *coalitionally acceptable* if for each history  $h$ , if  $S(h)$  is nonempty, then  $u_{S(h)}(x^\sigma(h)) \geq u_{S(h)}(x(h))$ . Finally, an absorbing process  $\sigma$  is *absolutely maximal* if at no history  $h$  does there exist a coalition  $T$  and a state  $y$  with  $T \in E(x(h), y)$ , such that  $u_T(x^\sigma(h, y, T)) \gg u_T(x^\sigma(h))$ . We discuss these concepts in more detail in Section 2.5.

**DEFINITION 1.** *A farsighted stable set  $F$  is absolutely maximal if it can be embedded in some absorbing, coalitionally acceptable, and absolutely maximal process  $\sigma$ ; that is,*

(i)  *$F$  is the set of all absorbing states of  $\sigma$ .*

(ii) *At any initial history  $h = \{x\}$  with  $x \notin F$ , or  $h = (x, (S, y))$  with  $x \in F$ ,  $S \in E(x, y)$  and  $y \notin F$ , the continuation chain from  $h$  is a blocking chain terminating in  $F$ .*

**2.5. Discussion.** Condition (i) asks that the set  $F$  be the ultimate repository of all end-states of  $\sigma$  starting from *any* history. That is, we seek not just absorption, but absorption back into  $F$ . Condition (ii) seeks consistency with the “blocking chain” approach that was originally used to describe  $F$ . That is, starting from some state not in  $F$ , or following some “challenge” to a state in  $F$  by another outside it, the process prescribes a blocking chain leading back into  $F$ , just as envisaged in the traditional definition of stability.

But, of course,  $\sigma$  does more: it prescribes continuation chains for *all* histories, not just those described in condition (ii) above. That gives us a setting where the counterfactual consequences of alternative actions can be discussed. We can consider deviations from ongoing chains, deviations from deviations, and so on;  $\sigma$  handles all these.

The requirement that  $\sigma$  be absolutely maximal is part of the embedding requirement for  $F$ . Note how that concept applies to every coalition, not just the coalition stipulated to move at the state in question: *no* coalition can stand to gain following *any* history. It is therefore stronger than the maximality condition of Dutta and Vohra (2017) which is imposed only on the coalition about to move, or their strong maximality condition, imposed only on coalitions that share a nonempty intersection with the coalition stipulated to move.<sup>8</sup> Absolute maximality is arguably the strongest form of maximality that one could insist on. These distinctions have bite, as illustrated by Example 1 below.

<sup>7</sup>I.e., there exists  $k$  with  $y^{(t)}(h) = x$  for  $t \geq k$ , where  $y^{(t)}$  is defined recursively in the obvious way.

<sup>8</sup>Absolute maximality is also stronger than the maximality conditions in Konishi and Ray (2003) and Ray and Vohra (2014). In a somewhat different context, this notion is also used by Xue (1998).

Our definition also asks that  $\sigma$  be absorbing: a negotiation process must ultimately terminate. Moreover, it asks for an absorbing state to be absorbing after *every* history leading to it.<sup>9</sup> This property does not follow from rationality per se; nothing dictates that a process *must* be absorbing: it could, for instance, cycle forever. We impose the condition as a desideratum of any negotiation process that “supports” the farsighted stable set.

In similar vein, coalitional acceptability is not a necessary concomitant of rationality, though sometimes it could be. It is a joint condition on any starting point and the final outcome. We view this property — that any coalition that moves at any stage must be made at least weakly better off in the final outcome — as a desirable characteristic of the negotiation process. A blocking chain satisfies coalitional acceptability, so for histories such as those described in Condition (ii), the latter imposes no additional restriction. Indeed, we could strengthen coalitional acceptability even more: we could ask that after *every* history ending in a state not in  $F$ , a blocking chain must be used, thereby imposing farsighted dominance not just “on path,” but following every conceivable history. We discuss this extension in Section 4.

### 3. THE MAXIMALITY OF FARSIGHTED STABLE SETS IN COALITIONAL GAMES

Our main theorem states that any farsighted stable set that satisfies two properties is absolutely maximal. In general, the direct construction of an absorbing, coalitionally acceptable and absolutely maximal process that embeds any given farsighted stable set — and thereby evaluating absolute maximality — is not an easy task. Our result is useful precisely because that task is replaced by the verification of two simple properties.

**3.1. Two Properties.** Consider the following two conditions:

**A.** *Suppose there are two states  $a$  and  $b$  in  $F$  such that  $u_j(b) > u_j(a)$  for some  $j$ . Then there exists a state  $z \in F$  such that  $u_j(z) \leq u_j(a)$ , and  $u_i(z) \geq u_i(b)$  for all  $i \neq j$ .*

**B.** *If  $a, b$  in  $F$ , there is no coalition  $T$  with  $u_T(b) \in V(T)$ ,  $T \in \pi(b)$  and  $u_T(b) \gg u_T(a)$ .*

Property A states that if player  $j$  gets a strictly higher payoff at  $b \in F$  than at  $a \in F$ , then it is possible to find another state in  $F$  at which  $j$ 's payoff is capped at  $u_j(a)$  without reducing the payoffs of the other players relative to those obtained under  $b$ . This property asks for some degree of transferability across payoffs in a farsighted stable set.

Property B states that given a state in  $F$ , there is no other state in  $F$  with a higher, feasible payoff for some coalition in that state. This property bears a close resemblance

---

<sup>9</sup>Dutta and Vartiainen (2018) consider a weaker notion of absorption. A “stable outcome” in their sense may be absorbing for some histories but not others. The set of such outcomes can be large. In a strictly superadditive game, the set of all strictly positive feasible payoffs is a “farsighted stable set” in their sense.

to internal stability. In fact, in the classical literature starting with von Neumann and Morgenstern (1944) and including Harsanyi (1974), it is internal stability.<sup>10</sup>

**3.2. Main Theorem and Discussion.** Our main result is

**THEOREM 1.** *If a farsighted stable set satisfies A and B, then it is absolutely maximal.*

Section 3.3 proves Theorem 1. Here, we discuss some aspects of this result.

**REMARK 1.** Both Properties A and B are satisfied by every farsighted stable set with a unique payoff profile.

This observation is immediate: with just one payoff profile in the stable set, the starting conditions in Properties A and B never occur, and so the properties are trivially valid. Dutta and Vohra (2017, Theorem 1) directly verify that, in fact, every single-payoff farsighted stable set satisfies maximality via a history-*independent* process.

Ray and Vohra (2015) characterize single-payoff farsighted stable sets using *separable* payoff allocations. Let  $u$  be an efficient payoff allocation; i.e., there is a state  $x$  with  $u(x) = u$  and no state  $x'$  with  $u(x') > u(x)$ . Allocation  $u$  is *separable* if whenever  $u_{S_i} \in V(S_i)$  for some pairwise disjoint collection of coalitions  $\{S_i\}$  that do not fully cover  $N$ , then  $u_T \in V(T)$  for some  $T \subseteq N - \cup_i S_i$ . For a feasible payoff profile  $u$ , let  $[u]$  be the collection of all states  $x$  such that  $u(x) = u$ . Ray and Vohra (2015, Theorem 2) show that  $[u]$  is a single payoff farsighted stable set if and only if  $u$  is separable.

The interior of the core is contained in the set of separable allocations, which are in turn contained in the coalition structure core. Every game for which the core has nonempty interior therefore possesses a single payoff farsighted stable set. Other known special cases in which some core allocation is separable are hedonic games with strict preferences and the top coalition property (Diamantoudi and Xue 2003), and matching games with strict preferences (Mauleon, Vannetelbosch and Vergote 2011).

We shall now make an important addition to this list. Among the most fruitful economic applications of coalitional games are those related to exchange economies. It is therefore of some significance that we can provide reasonable, sufficient conditions for a competitive equilibrium to yield a separable payoff allocation.

---

<sup>10</sup>In that literature, a coalition can move to any state as long the payoff restricted to the coalition is feasible for it; there is no restriction on the payoffs to outsiders. There, Property B is equivalent to internal (myopic) stability and is automatically satisfied by every stable set, farsighted or not. It is only because of our insistence on the coalitional sovereignty conditions (E.1) and (E.2) that Property B could go beyond internal stability, and therefore must be separately stated. In our setting, if there are  $a, b \in F$  and  $T$  such that  $u_T(b) \in V(T)$  and  $u_T(b) \gg u_T(a)$ , then by (E.2),  $T$  can move to some state, say  $b'$ , where  $u_T(b') = u_T(b)$ . But  $b'$  may not be in  $F$ , and while  $b$  is in  $F$ , it is also possible that  $T \notin E(a, b)$ , because the coalition structure and/or the payoffs of players outside  $T$  might differ across  $b$  and  $b'$ .



An exchange economy with a finite set of consumers  $N$  is denoted  $(N, \{X_i, u_i, \omega_i\}_{i \in N})$ , where  $X_i \subseteq \mathbb{R}^l$  is  $i$ 's consumption set,  $u_i : X_i \rightarrow \mathbb{R}$  is  $i$ 's utility function and  $\omega_i \in X_i$  is  $i$ 's initial endowment. A *competitive equilibrium* consists of  $(\{\xi_i\}, p)$ , where  $\xi_i$  denotes  $i$ 's consumption bundle, and  $p \in \mathbb{R}_+^l$  the vector of market prices, such that

- (i) for all  $i$ ,  $p \cdot \xi_i \leq p \cdot \omega_i$  and  $u_i(\xi'_i) > u_i(\xi_i)$  implies that  $p \cdot \xi'_i > p \cdot \omega_i$ , and
- (ii)  $\sum_{i \in N} \xi_i = \sum_{i \in N} \omega_i$ .

There is a natural way of constructing a coalitional game from a private ownership exchange economy. For every coalition  $S$ , let

$$V(S) = \{u_S \in \mathbb{R}^S \mid \exists \{\xi_i\}_{i \in S} \in \prod_{i \in S} X_i, \sum_{i \in S} \xi_i = \sum_{i \in S} \omega_i \text{ and } u_i(\xi_i) \geq u_i \text{ for all } i \in S\}.$$

REMARK 2. Assume that preferences are (a) locally non-satiated: for every  $\xi_i \in X_i$  there exists  $\xi'_i \in X_i$  arbitrarily close to  $\xi_i$  with  $u(\xi'_i) > u_i(\xi_i)$ ; and (b) strictly convex: if  $u_i(\xi'_i) \geq u_i(\xi_i)$  and  $\xi'_i \neq \xi_i$ , then  $u_i(t\xi'_i + (1-t)\xi_i) > u_i(\xi_i)$  for all  $t \in (0, 1)$ . Then the payoff profile  $u$  of any competitive equilibrium of an exchange economy is separable.

By Ray and Vohra (2015),  $[u]$  is a single-payoff farsighted stable set, and Remark 1 applies. For the proof and more discussion of Remark 2, see Online Appendix A.1.

We now remark on farsighted stable sets with nonsingleton payoffs. We do so for *simple games*, which are “transferable-utility” games with either “winning” or “losing” coalitions. Moreover, if a coalition is winning, then its complement is losing. A winning coalition has unit value that it can allocate among its players; a losing coalition has zero value. These games describe a rich class of situations: parliaments, bargaining institutions, and committees have been studied with this device.<sup>11</sup> In such games a state,  $x$ , can be described by its winning coalition  $W(x)$  (if any) and the payoff allocation  $u(x)$  among members of  $W(x)$ ; it is understood that  $u_i(x) = 0$  for all  $i \notin W(x)$ .

A *veto coalition* is a coalition with a losing complement, and a singleton veto coalition is a veto player. If the set of all veto players is winning, say that the game is *oligarchic*. Say that a non-oligarchic game is *standard* if it has a minimal veto coalition with replaceable players: that is, any of its members can be replaced by anyone else, and the resulting coalition would remain veto.

REMARK 3. Every farsighted stable set in a simple game satisfies Property B. Moreover, in every oligarchic or standard non-oligarchic simple game, there exists a farsighted stable set satisfying Property A, which is therefore absolutely maximal.

Online Appendix A.2 discusses related observations, and formally proves Remark 3.

<sup>11</sup>See Shapley 1962 for an introduction to simple games. Such games have been extensively analyzed in the context of the vNM stable set (see, e.g., Lucas 1992), are used in theories of bargaining (Baron and Ferejohn 1989) and have played a significant role in the analysis of political institutions; see, e.g., Winter (1996) and Austen-Smith and Banks (1999).

**3.3. Proof of Theorem 1.** We begin with an informal outline of the argument. Fix a farsighted stable set  $F$ . We want to stitch a collection of (coalitionally acceptable) chains together so that (a) these all terminate in  $F$ , and (b) whenever there is a deviation by *any* coalition at *any* stage of a chain, another (coalitionally acceptable) chain will start up that not only terminates in  $F$ , but also deters the deviation. Finally, (c) we would like to do this in a way so that starting from a particular state with no anterior history, or from a history following a single move away from  $F$ , we deploy a full-fledged blocking chain that makes every participating coalition strictly better off at its terminal node.

With these considerations in mind, we begin by fixing arbitrary blocking chains either from initial states with no history or following a single-step move from  $F$ . We pick these chains in any way we please from the collection of chains that already surround  $F$ : remain in  $F$  if already in  $F$ , or pick any blocking chain if starting from a state not in  $F$ . Think of these as describing “on-path” play. The proof consists of describing “off-path” play from every other history, constructed to deter all deviations, including deviations from off-path play. The construction of these paths will require us to invoke Properties A and B. The formal details of the proof assure us that (a) under these properties, such paths can be constructed, (b) that these paths themselves are coalitionally acceptable — every coalition initiating the paths will (weakly) benefit at journey’s end, and (c) that all paths, on-path or off-path, lead back to  $F$ . Now for a more formal account.

We first show that whenever there is a blocking chain from  $x$  to  $y$ , there exists what might be called a *canonical blocking chain* from  $x$  to  $y$ , in which each individual moves at most twice, possibly once at an intermediate step, and then again at the very last step, when “consolidation” occurs to generate the final state  $y$ .<sup>12</sup>

**LEMMA 1.** *Suppose that  $y$  farsightedly dominates  $x$  via the chain  $\{\tilde{y}^0, \tilde{y}^1, \dots, \tilde{y}^{\tilde{m}-1}, \tilde{y}^{\tilde{m}}\}$ ,  $\{\tilde{S}^1, \dots, \tilde{S}^{\tilde{m}}\}$ , where  $\tilde{y}^0 = x$  and  $\tilde{y}^{\tilde{m}} = y$ . Then there exists another (canonical) blocking chain  $\{y^0, y^1, \dots, y^{m-1}, y^m\}$ ,  $\{S^1, \dots, S^m\}$ , such that*

- (i)  $y^0 = x$  and  $y^m = y$ ; and
- (ii)  $S^i$  and  $S^j$  are disjoint for all  $i, j = 1, \dots, m - 1$ , where  $i \neq j$ .
- (iii)  $\cup_{k=1}^{m-1} S^k \subseteq S^m$ , so the set of all active movers in the canonical chain is  $S^m$ .

*Proof.* If  $\tilde{m} = 1$ , the original blocking chain is trivially a canonical blocking chain. So suppose  $\tilde{m} \geq 2$ . Set  $y^0 = x$  and  $S^1 = \tilde{S}^1$  and, if  $\tilde{m} > 2$ , then recursively let  $S^k = \tilde{S}^k - \cup_{t < k} \tilde{S}^t$  for all  $k = 2, \dots, \tilde{m} - 1$ . For any  $k = 1, \dots, \tilde{m} - 1$ , when coalition  $S^k$  moves, it does so by breaking into singletons. So, for any such  $k$ , the corresponding coalition structure,  $\pi^k$ , is such that all players in  $\cup_{t < k} \tilde{S}^t$  are in singletons, and (by Condition E.1)

<sup>12</sup>This is a generalization of the observation (Ray and Vohra (2015), Lemma 1) that in a simple game every blocking chain can be replaced by one with at most two steps. Either there is a one-step move by a winning coalition or there is an initial move by a veto coalition to force all players to a zero payoff state, followed by a final move by a winning coalition.

all other players belong to the same coalition as in  $\tilde{y}^k$ . At the last step, let  $S^{\tilde{m}} = \cup_{k=1}^{\tilde{m}} \tilde{S}^k$ . That is, we collect all the coalitions have already moved, along with all other individuals (if any) in  $\tilde{S}^{\tilde{m}}$ . Since  $S^{\tilde{m}}$  is the set of *all* players who were involved in moving from  $x$  to  $y$ , it is clearly effective in moving to  $y$ . Have it do so, creating the final coalition structure,  $\pi^{\tilde{m}} = \pi(y)$ .

Denote by  $u^k$  the associated payoffs in the newly constructed chain and by  $\tilde{u}^k = u(\tilde{y}^k)$  the payoffs generated by the original blocking coalition. Of course,  $u^0 = \tilde{u}^0 = u(x)$ , and  $u^{\tilde{m}} = \tilde{u}^{\tilde{m}} = u(y)$ . Given the coalition structures  $\pi^k, k = 1, \dots, \tilde{m}$  in the new chain, it follows from Conditions E.1 and E.2 that

$$(1) \quad \text{for } k = 1, \dots, \tilde{m} - 1, \quad u_i^k = 0 \text{ if } i \in \cup_{t \leq k} \tilde{S}^t, \text{ and } u_i^k = \tilde{u}_i^k \text{ otherwise.}$$

Let the associated states be  $y^k = (u^k, \pi^k)$  for all  $k = 1, \dots, \tilde{m} - 1$ , and  $y^{\tilde{m}} = y$ . It is possible that for some stages  $k < \tilde{m}$ ,  $S^k$  as defined is empty and the succeeding state  $y^{k+1}$  is identical to  $y^k$ .<sup>13</sup> In that case, remove the step at all such  $k$ . We are left with a chain of  $m$  steps, where  $m \leq \tilde{m}$ , and this is the chain to which the lemma refers. By construction, (i) and (ii) are satisfied. Because  $S^m = S^{\tilde{m}}$  is the set of all players involved in the move,  $\cup_{k=1}^{m-1} S^k$  is contained in  $S^m$  and (iii) is also satisfied.

We only need to check that the new chain is a blocking chain. That is, for every  $k \geq 1$  and every  $i \in S^k$ ,  $u_i(y) > u_i^{k-1}$ . But this is true because  $u_i(y) > \tilde{u}_i^{k-1}$  since the original chain is a blocking chain, and by (1),  $\tilde{u}_i^{k-1} \geq u_i^{k-1}$ . ■

Consider any farsighted stable set  $F$ . For each  $x \notin F$ , fix any blocking chain,  $\mathbf{c}(x)$ , and define  $\Psi(x) \in F$  to be its terminal state. If  $x \in F$ , define  $\Psi(x) = x$ . The next lemma uses Properties A and B and the existence of a canonical blocking chain to construct a particular chain that will be later used to deter deviations from some on-path process.

**LEMMA 2.** *Let a farsighted stable set  $F$  satisfy Properties A and B. Consider states  $x$  and  $y$  with  $x \notin F$ ,  $\Psi(x) = a$  and  $\Psi(y) = b$ . For any  $T \in E(x, y)$ , there is  $z \in F$  and a coalitionally acceptable chain from  $y$  to  $z$  with  $u_j(z) \leq u_j(a)$  for some  $j \in T$ .*

*Proof.* Fix states  $x, y, a, b$  and a coalition  $T$  as in the statement of the lemma. Because any nonempty blocking chain is acceptable, there is nothing to prove if  $u_j(b) \leq u_j(a)$  for some  $j \in T$ ; simply take  $z = b$  and use the original chain from  $y$  to  $b$ . On the other hand, if  $u_T(b) \gg u_T(a)$ , then by Property B, no subset of  $T$  belongs to the coalition structure at  $b$ . Therefore  $y \neq b$ , so that  $y \notin F$ . By Lemma 1, there is a canonical blocking chain from  $y$  to  $b$ . Fix one such canonical blocking chain,  $\mathbf{c} = \{y, y^1, \dots, y^{m-1}, y^m\}, \{S^1, \dots, S^m\}$ , where  $(y^m, S^m) = b$ . Recall that  $S^i$  and  $S^j$  are disjoint for all  $i, j = 1, \dots, m - 1$  with  $i \neq j$ , and  $\cup_{k=1}^{m-1} S^k \subseteq S^m$  is the set of all active movers in the blocking chain. Since no subset of  $T$  belongs to the partition at  $b$ , every player in  $T$  is involved in some coalitional move in this blocking chain; that is,  $T \subset S^m$ .

<sup>13</sup>This will happen when a new coalition belongs to the union of previous coalitions in the chain.

We now consider two cases:

*Case 1.* Some subset  $T'$  of  $T$  moves *only* in the final step from  $y^{m-1}$  to  $b$ , and so is part of the coalition  $W \equiv S^m - \cup_{k=1}^{m-1} S^k$ . Pick any  $j \in T'$ . Modify the original chain by adding an extra step after  $y^{m-1}$  in which  $W - j$  breaks up into singletons and moves from  $y^{m-1}$  to  $y'$ . At  $y'$  all players in  $S^m$  are in singletons and the subpartition of the remaining players is the same as it is at  $b$ . By Condition E.1, the latter property implies:

$$(2) \quad u_k(y') = u_k(b) \text{ for all } k \in N - S^m.$$

With  $y'$  inserted between  $y^{m-1}$  and  $y^m$  we have a chain  $c' = \{y, y^1, \dots, y^{m-1}, y', y^m\}$ ,  $\{S^1, \dots, S^{m-1}, W - j, S^m\}$ , with  $(y^m, S^m) = b$ .<sup>14</sup> Clearly, this new chain is also a blocking chain. There are two critical features of this new blocking chain: (1) at state  $y'$ ,  $u_j(y') = 0$  and player  $j$  has yet to move ; (2)  $u_i(y') = u_i(b)$  for all  $i \in N - S^m$ .

Property A assures us of the existence of  $z \in F$  such that  $u_j(z) \leq u_j(a)$  and  $u_i(z) \geq u_i(b)$  for all  $i \neq j$ . Modify the blocking chain  $c'$  by replacing the terminal state with  $z$  to construct the chain  $\bar{c} = \{y, y^1, \dots, y^{m-1}, y', z\}$ ,  $\{S^1, \dots, S^{m-1}, W - j, S'\}$ , where  $S'$  is a minimal set of players needed to replace  $y'$  with  $z$ . We claim that  $\bar{c}$  is a coalitionally acceptable chain. Note that in addition to  $S^m - j$ , it is possible that other active players appear in  $S'$ , including player  $j$ . Consider the acceptability condition for each of these sets of players in turn.

(i) For every  $k = 1, \dots, m - 1$  and any  $i \in S^k$ ,  $u_i(b) > u_i^{k-1}$  because  $c$  is a blocking chain. Since  $u_i(z) \geq u_i(b)$  for any such  $i$  we also have  $u_i(z) > u_i^{k-1}$ . Since  $u_i(y') = 0$  for all such  $i$ ,  $u_i(z) > u_i(y')$  as well. (All such players must therefore belong to  $S'$ ). Exactly the same argument applies to all players in  $W - j$ . Thus all players in  $S^m - j$  strictly gain by participating in  $\bar{c}$ .

(ii) Any player  $k \in S' - S^m$  moves only at the last step, from  $y'$  to  $z$ . From (2) we know that  $u_k(y') = u_k(b)$ , so  $u_k(z) \geq u_k(y')$ . Thus any such player is (weakly) willing to participate in  $\bar{c}$ .

(iii) Suppose  $j \in S'$ . Recall that  $c'$  was constructed to make sure that at state  $y'$ ,  $j$  has yet to move and  $u_j(y') = 0$ . This means that  $u_j(z) \geq u_j(y')$  and player  $j$  is also (weakly) willing to participate.

These three steps together prove that  $\bar{c}$  is a coalitionally acceptable chain.<sup>15</sup>

*Case 2.*  $T \subseteq \cup_{t < m} S^t$ ; i.e., every member of  $T$  has made some move by the time the state  $y^{m-1}$  is reached. Let  $k < m$  be the maximal index such that some member of  $T$  belongs to  $S^k$ , and let  $j$  be any such member of  $T$ . Since all players in  $T - S^k$  have

<sup>14</sup>If  $W = \{j\}$  this step is redundant:  $y' = y^{m-1}$ . However, it is still the case that at  $y'$  player  $j$  has yet to move and  $u_j(y') = 0$ .

<sup>15</sup>There are precisely two reasons why  $\bar{c}$  may not be a blocking chain, as captured through cases (ii) and (iii) above: (a) there exists a player  $k \in S' - S^m$  with  $u_k(z) = u_k(y') = u_k(b)$ , or (b) Player  $j$  is included in  $S'$  and  $u_j(y') = u_j(z) = 0$ . We will make use of this observation in Section 4.

already moved, a move by  $S^k$  must mean that  $S^k$  is not a singleton, i.e.,  $S^k - j \neq \emptyset$ . Now interpret the move to  $y^k$  as one made by  $S^k - j$ . Because  $S^k$  breaks into singletons in the canonical chain, this interpretation is valid.<sup>16</sup> Keep the rest of the process unchanged until  $y^{m-1}$ . With this interpretation we have a blocking chain in which there is a player  $j \in T$  who at state  $y^{m-1}$  has yet to make a move. In other words, there is a subset of  $T$  that moves *only* in the final step from  $y^{m-1}$  to  $y^m$ . But then we are back in Case 1. ■

Recall that for a given farsighted stable set we have chosen for every  $x \notin F$  some blocking chain  $\mathbf{c}(x)$  with terminal state  $\Psi(x)$ . We will now embed this in an absorbing process  $\sigma$  satisfying absolute maximality and coalitional acceptability. Recall that for any history  $h$ ,  $x(h)$  denotes the current state. Let  $\ell(h)$  denote the state immediately preceding  $x(h)$ , in case there is one. In what follows, we will recursively assign, not just  $\sigma$ , but an entire chain  $\mathbf{c}(h)$  following each history  $h$ , taking care to “follow through” with appropriate continuations for nested collections of histories.

For any history  $h$  with current state  $x(h) \in F$ , let  $\sigma$  prescribe no change, i.e., if  $x(h) \in F$ ,  $\sigma(h) = (x(h), \emptyset)$ . Now consider histories in which the current state is not in  $F$ .

Begin with a single-state history, or a one-step history,  $h = \{x\}$  (where  $x \notin F$ ). Set  $\mathbf{c}(h) = \mathbf{c}(x)$ , the already-fixed blocking chain that leads from  $x$  to terminal state  $\Psi(x)$ . The associated  $\sigma$  is given by  $\sigma(h) = (y(h), S(h))$ , which picks up the initial step in  $\mathbf{c}(x)$ . (When the definition is complete, we will also see that  $x^\sigma(h) = \Psi(x)$ .)

Next, consider any history  $h$  such that  $x(h) \notin F$ , but  $\ell(h) \in F$ . In this case, let  $\sigma$  specify exactly the same move as in the previous paragraph starting from  $x = x(h)$ , so that  $\mathbf{c}(h) = \mathbf{c}(x(h))$ , with the associated  $\sigma(h)$  defined accordingly.

It remains to define the process for histories of the form  $h$  where  $x(h) \notin F$  and  $\ell(h) \notin F$ . Recursively, suppose that we have attached a chain  $\mathbf{c}(h)$  to every history  $h$  with  $K$  steps or less, where  $K \geq 1$ . Now consider a history  $h$  with  $K + 1$  steps. Let  $h_K$  denote the first  $K$  steps. There are now three possibilities:

(i) If  $x(h) = y(h_K)$ , where  $y(h_K)$  is specified by  $\mathbf{c}(h_K)$ , then simply use the continuation chain of  $\mathbf{c}(h_K)$  at  $h$ , and define  $\sigma(h)$  accordingly.

(ii) If  $x(h) = \ell(h) \neq y(h_K)$ , restart  $\mathbf{c}(h_K)$ : set  $\mathbf{c}(h) = \mathbf{c}(h_K)$  and  $\sigma(h) = \sigma(h_K)$ .

(iii) If  $x(h) \neq y(h_K)$  and  $x(h) \neq \ell(h)$ , let  $T$  be the associated coalition in the last step of the history  $h$ , to be interpreted as the coalition that “deviated” from  $x(h_K)$  to  $x(h)$ , instead of the prescribed move to  $y(h_K)$ . Let  $a$  equal the “intended” terminal state from  $h_K$  (under  $\mathbf{c}(h_K)$ ), and let  $y$  equal  $x(h)$ . By Lemma 2, there is a state  $z \in F$  and an acceptable chain  $\mathbf{c}'$  from  $y$  to  $z$  such that  $u_j(z) \leq u_j(a)$  for some  $j \in T$ . Fix any such chain  $\mathbf{c}'$  and assign it to the history  $h$ , defining  $\sigma$  accordingly at  $h$ . This last step ensures that for no  $h$  can a coalition profitably deviate from the path prescribed by  $\mathbf{c}(h)$ .

<sup>16</sup>Formally, replace  $S^k$  by  $S^k - j$ .

Proceeding recursively in this way, we define  $c(h)$  for every  $h$ , along with the accompanying  $\sigma(h)$ . Clearly,  $\sigma$  embeds  $F$  and is coalitionally acceptable.<sup>17</sup> For a history  $h$  with  $x(h) \in F$ , absolute maximality follows from the farsighted internal stability of  $F$ . If  $x(h) \notin F$ , absolute maximality follows from the last step of the previous paragraph. ■

**3.4. The Tightness and Necessity of Properties A and B.** As discussed, there is a sizable class of games with farsighted stable sets satisfying Properties A and B. Can these properties be dispensed with for free? The answer is no.

**EXAMPLE 1 (Tightness of Property A).** We exhibit a farsighted stable set that fails Property A, satisfies Property B, and is not absolutely maximal.

Consider a four-player simple game in which a coalition is winning if and only if it weakly contains one of these minimal winning coalitions:  $\{1, 2, 3\}$ ,  $\{1, 4\}$ ,  $\{2, 4\}$  and  $\{3, 4\}$ . Let  $m = (1/3, 1/3, 1/3, 2/3)$ . For every minimal winning  $S$ , define the profile  $u^S$  by  $u_i^S = m_i$  for  $i \in S$  and  $u_i^S = 0$  for  $i \notin S$ . Let  $F$  be the farsighted stable set corresponding to the collection of all such utility profiles —  $(1/3, 1/3, 1/3, 0)$ ,  $(1/3, 0, 0, 2/3)$ ,  $(0, 1/3, 0, 2/3)$  and  $(0, 0, 1/3, 2/3)$  — along with the respective winning coalitions. By Remark 3,  $F$  satisfies Property B. But it does not satisfy Property A. To see this, consider the states  $a, b \in F$  where  $u(a) = (1/3, 0, 0, 2/3)$  and  $u(b) = (1/3, 1/3, 1/3, 0)$ . There is no  $z \in F$  with  $u_3(z) = 0$ ,  $u_1(z) \geq 1/3$ ,  $u_2(z) \geq 1/3$ . So Property A fails, and we cannot appeal to Theorem 1 to show that  $F$  is absolutely maximal. In fact, Online Appendix A.5 shows that under some restrictions on the effectivity correspondence,  $F$  is not absolutely maximal. Thus, Property A cannot be freely removed from the statement of Theorem 1.

It is of interest to note that this is a regular non-oligarchic simple game, and so by Remark 3 it does have an absolutely maximal farsighted stable set; see Remark 4 in Appendix A.2 for details. Absolute maximality can therefore refine the collection of farsighted stable sets of a given game.

**EXAMPLE 2 (Tightness of Property B).** We exhibit a farsighted stable set that satisfies Property A, fails Property B, and is not absolutely maximal.

Consider a six-player game in which each coalition  $S$  has only one efficient payoff  $\nu(S)$ . Players 1 and 5 are symmetric, as are players 2 and 4. Player 3 gets a constant payoff whenever her payoff is positive. Player 6 always gets a zero payoff. Players 3 and 6 create synergies with other players. Player 3 generally benefits from those synergies

<sup>17</sup>In fact, barring case (iii), every  $h$  with  $x(h) \notin F$  is assigned to a *blocking* chain terminating in  $F$ .

herself; player 6 is completely indifferent throughout. Formally:

$$\begin{aligned} \nu(\{1, 2\}) &= \nu(\{4, 5\}) = (3, 3), & \nu(\{1, 3\}) &= \nu(\{3, 5\}) = (2, 2), \\ \nu(\{2, 3, 4\}) &= (4, 2, 4), & \nu(\{1, 3, 5\}) &= (1, 2, 1) \\ \nu(\{1, 3, 4, 5, 6\}) &= (3, 2, 4, 3, 0), & \nu(\{1, 2, 3, 5, 6\}) &= (3, 4, 2, 3, 0) \\ \nu(\{2, 3, 4, 5, 6\}) &= (4, 2, 4, 3, 0), & \nu(\{1, 2, 3, 4, 6\}) &= (3, 4, 2, 4, 0), \\ \nu(S) &= 0 \text{ for all other } S. \end{aligned}$$

There are as many states as there are coalition structures. However, many of them have the same payoff profile and differ only in the way in which some zero-payoff players are partitioned. To describe the collection of states that have the same payoff we need some more notation. For every coalition  $S$ , let  $\pi_S$  denote a subpartition of  $S$  and let  $\Pi^0(S) = \{\pi_S \mid \nu(T) = 0 \text{ for all } T \in \pi_S\}$  be the collection of subpartitions that result in every player in  $S$  getting 0.<sup>18</sup>

States	Structures	Payoffs to Players					
		1	2	3	4	5	6
$X^1$	$\{1, 2\}, \{3, 5\}, \Pi^0(\{4, 6\})$	3	3	2	0	2	0
$X^2$	$\{1, 3\}, \{4, 5\}, \Pi^0(\{2, 6\})$	2	0	2	3	3	0
$X^3$	$\{2, 3, 4\}, \Pi^0(\{1, 5, 6\})$	0	4	2	4	0	0
$x^4$	$\{1\}, \{2, 3, 4, 5, 6\}$	0	4	2	4	3	0
$x^5$	$\{1, 3, 4, 5, 6\}, \{2\}$	3	0	2	4	3	0
$x^6$	$\{1, 2, 3, 5, 6\}, \{4\}$	3	4	2	0	3	0
$x^7$	$\{1, 2, 3, 4, 6\}, \{5\}$	3	4	2	4	0	0

TABLE 1. Farsighted Stable Set for Example 2.

Consider the set of states  $F = X^1 \cup X^2 \cup X^3 \cup \{x^4, x^5, x^6, x^7\}$  shown in Table 1. It is easy to see that  $F$  satisfies Property A and fails Property B. To see the former, notice that  $x^4$  is a state at which Player 1 receives 0, her worst payoff. At that state each of the other players is getting their *maximum* possible payoff. We can make a parallel argument for Players 2, 4, and 5 (using states  $x^5, x^6$  and  $x^7$ , respectively).<sup>19</sup> Players 3 and 6 have payoffs that are invariant in  $F$ . So Property A is fully verified. However,  $F$  does not satisfy Property B. Coalition  $\{4, 5\}$  prefers a state in  $X^2$  to a state in  $X^1$  — the payoffs are  $(3, 3)$  in the former, compared to  $(2, 2)$  in the latter — and it can achieve the payoff  $(3, 3)$  on its own.

Appendix A.5 shows that  $F$  is a farsighted stable set, but it is not absolutely maximal. This shows that Property B cannot be dispensed with in our main theorem.

<sup>18</sup>For instance,  $\Pi^0(\{1, 2, 3\}) = \{\{\{1\}, \{2\}, \{3\}\}, \{\{1, 2, 3\}\}, \{\{1\}, \{2, 3\}\}\}$ .

<sup>19</sup>The only role for player 6 and of states  $x^4, x^5, x^6$  and  $x^7$  is to ensure that Property A is satisfied.

These examples also demonstrate that full history dependence (and zero discounting, as implicitly assumed) does not mean that anything goes. It is not the case that *any* farsighted stable set can be embedded in a coalitionally rational and absolutely maximal process. In short, a folk theorem is not to be had in the current context, particularly when we view the solution concept as pertaining to a *set* of states, which — in the spirit of von Neumann and Morgenstern stability — is the right thing to do.

At the same time, Properties A and B are not necessary for a farsighted stable set to be absolutely stable:

EXAMPLE 3. An absolutely farsighted stable set that does not satisfy Property A or Property B. Consider a five-player simplification of Example 2:

$$\begin{aligned} \nu(\{1, 2\}) &= \nu(\{4, 5\}) = (3, 3), & \nu(\{1, 3\}) &= \nu(\{3, 5\}) = (2, 2), \\ \nu(\{2, 3, 4\}) &= (4, 2, 4), & \nu(S) &= 0 \text{ for all other } S. \end{aligned}$$

Appendix A.5 shows that  $F = \{x^1, x^2\} \cup X^3$ , described in Table 2, is a farsighted stable set, *and* that it satisfies absolute maximality. Yet, Property B fails in this example for the same reason as in Example 2. Property A fails because player 1 prefers  $x^1$  to any state in  $X^3$  but there is no state  $x \in F$  such that  $u_1(x) = 0$  and  $u_5(x) \geq 2$ . So A and B are not logically necessary for absolute maximality of a farsighted stable set.

States	Structures	Payoffs to Players				
		1	2	3	4	5
$x^1$	$\{1, 2\}, \{3, 5\}, \{4\}$	3	3	2	0	2
$x^2$	$\{1, 3\}, \{4, 5\}, \{2\}$	2	0	2	3	3
$X^3$	$\{2, 3, 4\}, \Pi^0(\{1, 5\})$	0	4	2	4	0

TABLE 2. Farsighted Stable Set for Example 3.

#### 4. REMARKS ON COALITIONAL ACCEPTABILITY

We have embedded a farsighted stable set in an ambient process. The process satisfies maximality and coalitional acceptability starting from any history. Our embedding additionally asks for those paths that lead to the farsighted stable set from any initial *singleton* history, or following a single deviation from the stable set, to be blocking chains that generate *strict* gains to the moving coalitions. The idea is to achieve a complete embedding not just of the set, but of the usual blocking chains that are used to describe how “external stability” is achieved. In this section, we remark on the (weaker) coalitional acceptability requirement that we’ve placed on other chains following “off-path” histories.



To understand why a condition such as coalitional acceptability should be included, pick any state  $x$ , however unpalatable for some or all of the players, and then define a Markovian  $\sigma$  that specifies a chain back to  $x$  from every state. There is no way out of such an impasse: the process satisfies absolute maximality, but it is an absurd process nonetheless. A slight perturbation of the model will fix that. Suppose that each coalition receives payoffs in real time and discounts future payoffs. Then coalitional acceptability would be *implied* by the no-single-deviation condition, and it would remain in place as the discount factor converges to one, which is the model analyzed here.<sup>20</sup>

Indeed, we could strengthen coalitional acceptability even more: we could ask that after *every* history ending in a state not in  $F$ , a blocking chain must be used, thereby imposing farsighted dominance not just “on path,” but following every conceivable history.<sup>21</sup> To this end, say that  $\sigma$  is a *blocking process* if for each history  $h$ , if  $S(h)$  is nonempty and  $x(h) \neq x^\sigma(h)$ , then  $u_{S(h)}(x^\sigma(h)) > u_{S(h)}(x(h))$ .

Could we use a blocking process to embed a farsighted stable set? An inspection of the proof of Theorem 1 reveals that it will suffice to strengthen Lemma 2 so that the coalitionally acceptable chain constructed to deter deviations is in fact a blocking chain. At a minimum, this will require that when we dissuade an off-path deviation by finding a coalitionally acceptable chain from  $y$  to  $z \in F$ , all players involved in this chain must receive a strictly positive payoff at  $z$ . With this in mind, say that a state  $x$  is *regular* if  $u_i(x) > 0$  for every  $i$  such that  $i \in S \in \pi(x)$  and  $u_S(x) > 0$ .<sup>22</sup> We modify Property A to refer to regular states.

**Property A'.** *Suppose there are two regular states  $a$  and  $b$  in  $F$  such that  $u_j(b) > u_j(a)$  for some  $j$ . Then there exists a regular state  $z \in F$  such that  $u_j(z) \leq u_j(a)$ , and  $u_i(z) \geq u_i(b)$  for all  $i \neq j$ .*

**PROPOSITION 1.** *A farsighted stable set  $F$  can be embedded in an absorbing and absolutely maximal blocking process in any of the following circumstances:*

- (i)  $F$  is a single-payoff farsighted stable set.
- (ii)  $F$  satisfies property A' in a simple game.
- (iii)  $F$  satisfies property A' and for every  $y \notin F$  there is a blocking chain from  $y$  to  $x \in F$  with  $\pi(x) = N$ .

<sup>20</sup>That said, coalitional acceptability is not always implied by a discounted real-time model of negotiations. For instance, if a coalition refuses to move, then another coalition might be called upon to do so; and indeed it is possible to lock two coalitions into a coordination failure so that coalitional acceptability applies to neither of them.

<sup>21</sup>This has the virtue of generating symmetric requirements for on-path and off-path processes. On the other hand, one could argue — as is implicit in Definition 1 — that as long as we are incentivizing all coalitions to move, there is no reason to insist that they be *strictly* better off.

<sup>22</sup>In a simple game this reduces to  $u_i(x) > 0$  for all  $i \in W(x)$ , as in Ray and Vohra (2015).

Part (i) of the Proposition shows that Remark 1 can be strengthened to apply to embedding in a blocking process and (ii) shows that Remark 3 can be similarly strengthened because many simple games possess farsighted stable sets satisfying Property A'; see Remarks 4 and 5 in the Online Appendix.

*Proof.* (i) By Dutta and Vohra (2017), Theorem 1, a single-payoff farsighted stable set can be embedded in a history-independent (stationary) process that is absorbing, blocking and maximal. Clearly *no* coalition can find a profitable deviation when the final payoff is unique, so absolute maximality also holds. To prove the remaining two cases we rely on a modified version of Lemma 2, which is the Claim below. With this Claim in place of Lemma 2, the proof of the Proposition is the same as that of Theorem 1.

**Claim.** Consider a farsighted stable set  $F$  that satisfies Conditions (ii) or (iii) of Proposition 1. Suppose  $T$  moves from state  $x \notin F$  to state  $y$ ,  $\Psi(x) = a$  and  $\Psi(y) = b$ . Then there is a state  $z \in F$  and a blocking chain from  $y$  to  $z$  such that  $u_j(z) \leq u_j(a)$  for some  $j \in T$ .

Consider states  $x, y, a, b$  and a coalition  $T$  as in the statement of the Claim. If the conclusion of the Claim is false,  $u_T(b) \gg u_T(a)$ . Conditions (ii) and (iii) both imply that  $F$  satisfies Property B. Thus, in either case,  $y \neq b$ , so that  $y \notin F$ . Proceeding as in the proof of Lemma 2, we have a player  $j \in T$  and a blocking chain  $\mathbf{c}' = \{y, y^1, \dots, y^{m-1}, y', y^m\}$ ,  $\{S^1, \dots, S^{m-1}, W - j, S^m\}$  where  $y^m = b$  and  $S^m = \bigcup_{j=1}^{m-1} S^j \cup W$  is the set of all players who actively move in the blocking chain  $\mathbf{c}'$ . By Condition A', there is a regular state  $z$  such that  $u_j(z) \leq u_j(b)$  and  $u_i(z) \geq u_i(b)$  for all  $i \neq j$ . Modify  $\mathbf{c}'$  by replacing the terminal state with  $z$  to construct the chain  $\bar{\mathbf{c}} = \{y, y^1, \dots, y^{m-1}, y', z\}$ ,  $\{S^1, \dots, S^{m-1}, W - j, S'\}$ , where  $S'$  is a minimal set of players needed to replace  $y'$  with  $z$ . We will show that  $\bar{\mathbf{c}}$  is a blocking chain. As in the proof of Lemma 2, we know that  $S'$  includes all players in  $S^m - j$ , all of whom strictly gain by participating in  $\bar{\mathbf{c}}$ . As pointed out in Footnote 15, there can be only two possible reasons why  $\bar{\mathbf{c}}$  may be a coalitionally acceptable but not a blocking chain:

- (a) there exists a player  $k \in S' - S^m$  for whom  $u_k(z) = u_k(y') = u_k(b)$ , or
- (b)  $j \in S'$  and  $u_j(z) = u_j(y') = 0$ .

To complete the proof we will rule out each of these possibilities.

Suppose (a) holds. If the game is a simple game,  $S^m$  is the winning coalition at  $b$  and  $k \notin S^m$  implies that  $u_k(b) = 0$ . Since  $k \in S'$  and  $z$  is a regular state,  $u_k(z) > 0$ , so we cannot have  $u_k(z) = u_k(b)$ . Under Condition (ii) of the Proposition, without loss of generality,  $\pi(b) = N$  so  $S^m = N$ , which implies that  $S' - S^m = \emptyset$  and again (a) is not possible.

Suppose  $j \in S'$ . Since  $z$  is a regular state,  $u_j(z) > 0$ , so (b) cannot hold. ■

To go beyond the circumstances described in Proposition 1, in addition to Property B, a strengthening of Property A' will be needed to handle possibility (a) in the proof of the Claim; see the Online Appendix for details.

## REFERENCES

- Aumann, Robert and Michael Maschler (1964), "The bargaining set for cooperative games", in *Advances in Game Theory* (M. Dresher, L. S. Shapley, and A. W. Tucker, Eds.), *Annals of Mathematical Studies* No. 52, Princeton Univ. Press, Princeton, NJ.
- Aumann, Robert and Roger Myerson (1988), "Endogenous Formation of Links Between Players and of Coalitions, An Application of the Shapley Value," in *The Shapley Value: Essays in Honor of Lloyd Shapley*, Alvin Roth, ed., 175–191. Cambridge: Cambridge University Press.
- Austen-Smith, David and Jeffrey Banks (1999), *Positive Political Theory I: Collective Preference*, Ann Arbor, MI: University of Michigan Press.
- Baron, David and John Ferejohn (1989), "Bargaining in Legislatures," *American Political Science Review*, **83**, 1181–1206.
- Bloch, Francis (1996), "Sequential Formation of Coalitions in Games with Externalities and Fixed Payoff Division," *Games and Economic Behavior* **14**, 90–123.
- Bloch, Francis and Anne van den Nouweland (2017), "Farsighted stability with heterogeneous expectations", mimeo.
- Chwe, Michael (1994), "Farsighted Coalitional Stability," *Journal of Economic Theory*, **63**, 299–325.
- Diamantoudi, Effrosyni and Licun Xue (2003), "Farsighted Stability in Hedonic Games," *Social Choice and Welfare* **21**, 39–61.
- Dutta, Bhaskar and Rajiv Vohra (2017), "Rational Expectations and Farsighted Stability", *Theoretical Economics* **12**, 1191–1227.
- Dutta, Bhaskar and Hannu Vartiainen (2018), "Coalition Formation and History Dependence", mimeo.
- Chatterjee, Kalyan, Bhaskar Dutta, Debraj Ray and Kunal Sengupta (1993), "A Non-cooperative Theory of Coalitional Bargaining," *The Review of Economic Studies* **60**, 463–477.
- Greenberg, Joseph (1990), *The Theory of Social Situations: An Alternative Game-Theoretic Approach*, Cambridge, UK: Cambridge University Press.
- Herings, P. Jean-Jacques, Ana Mauleon, and Vincent Vannetelbosch (2004), "Rationalizability for social environments", *Games and Economic Behavior* **49**, 135–156.
- (2017), "Matching with Myopic and Farsighted Players", mimeo.
- Harsanyi, John (1974), "An Equilibrium-Point Interpretation of Stable Sets and a Proposed Alternative Definition," *Management Science* **20**, 1472–1495.

- Jordan, James (2006), “Pillage and Property,” *Journal of Economic Theory* **131**, 26–44.
- Kimya, Mert (2015), “Equilibrium Coalitional Behavior”, mimeo, Brown University.
- Konishi, Hideo and Debraj Ray (2003), “Coalition Formation as a Dynamic Process,” *Journal of Economic Theory* **110**, 1–41.
- Lucas, William (1992), “von Neumann-Morgenstern Stable Sets,” in *Handbook of Game Theory, Volume 1*, ed. by Robert Aumann, and Sergiu Hart, 543–590. North Holland: Elsevier.
- Mauleon, Ana, Vincent Vannetelbosch and Wouter Vergote (2011), “von Neumann-Morgenstern farsighted stable sets in two-sided matching,” *Theoretical Economics* **6**, 499–521.
- Okada, Akira (1996), “A Noncooperative Coalitional Bargaining Game with Random Proposers,” *Games and Economic Behavior* **16**, 97–108.
- Ray, Debraj (2007), *A Game-Theoretic Perspective on Coalition Formation*, Oxford University Press.
- Ray, Debraj and Rajiv Vohra (1999), “A Theory of Endogenous Coalition Structures,” *Games and Economic Behavior* **26**, 286–336.
- (2014), “Coalition Formation,” in *Handbook of Game Theory, Volume 4*, ed. by H. Peyton Young and Shmuel Zamir, 239–326. North Holland: Elsevier.
- (2015), “The Farsighted Stable Set”, *Econometrica*, **83**, 977–1011.
- Ray, Debraj (©) Arthur Robson (2018), “Certified Random: A New Order for Coauthorship,” *American Economic Review* **108**, 489–520.
- Shapley, Lloyd (1962), “Simple Games: An Outline of the Descriptive Theory,” *Behavioral Science* **7**, 59–66.
- Vohra, Rajiv (©) Debraj Ray (2018), “Maximality in the Farsighted Stable Set,” Brown University, Department of Economics Working Paper 2017-4, September 2018.
- von Neumann, John and Oskar Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton, NJ: Princeton University Press.
- Winter, Eyal (1996), “Voting and Vetoing,” *American Political Science Review* **90**, 813–823.
- Xue, Licun (1998), “Coalitional Stability under Perfect Foresight,” *Economic Theory* **11**, 603–627.