MAXIMALITY IN THE FARSIGHTED STABLE SET

Rajiv Vohra ⓡ Debraj Ray[†]

September 2018

Abstract. The stable set of von Neumann and Morgenstern imposes credibility on coalitional deviations. Their credibility notion can be extended to cover farsighted coalitional deviations, as proposed by Harsanyi (1974), and more recently reformulated by Ray and Vohra (2015). However, the resulting farsighted stable set suffers from a conceptual drawback: while coalitional deviations 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 this stringent maximality requirement. These properties are satisfied by many, but not all, known farsighted stable sets.

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

JEL CLASSIFICATION: C71, D72, D74

1. INTRODUCTION

The core is a classical solution concept: it identifies payoff profiles that no group, or coalition, can dominate with an allocation that is feasible for the coalition in question. But the core does not ask if the new allocation itself is threatened or "blocked" by *other* coalitions. In this conceptual sense the solution is too strong, possibly excluding allocations that would not be credibly dominated. The problem is that the definition of credibility is often circular — an allocation is not credible if it is not challenged by a credible allocation. Concepts such as the bargaining set (Aumann and Maschler 1964), which only try to build in an additional "round" of domination, are just not up to the task. But the vNM stable set (von Neumann and Morgenstern, 1944) can indeed be seen as such a theory: it 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 $u' \in Z$.

[†]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. Names are in random order, as proposed in Ray ⓡ Robson (2018). We are grateful to three anonymous referees and a Co-Editor for many helpful comments.

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

Notice how internal and external stability work in tandem to get around the circularity implicit in the definition of credibility. The set Z is to be viewed as a "standard of behavior" (Greenberg, 1990). Once accepted, no allocation in the standard can be overturned by another allocation also satisfying the standard. Moreover, allocations within the standard jointly dominate all non-standard allocations. This perspective drives home the idea that the relevant solution concept is not *a* payoff profile, but a *set* of payoff profiles which 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.[1]

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 be better off (relative to their starting points) once the process has come to a standstill. 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 of concern not just at any stage but 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,

[1]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), 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).

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 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 $u(a) = (1, 1)$, $u(b) = (0, 0)$, $u(c) = (10, 10)$ and $u(d) = (0, 20)$. Suppose that state a can only 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," assigning a particular player or coalition to move at each node. Noncooperative game theory imposes such protocols,[2] but the apparent gain in precision is in part illusory, for it is well known that 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. Specifically, this implies that the problem of maximality is not just restricted to the coalition that actually moves, but it also applies to other coalitions that could *potentially* move. So, while maximality is certainly related to sequential rationality or subgame perfection, it goes beyond that. That is, different definitions of maximality are possible depending on which coalitions are "allowed" to move at any state.

The weakest of these conditions, 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.*

[2]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).

None of this is particularly germane to the example above: all the concepts above coincide, and the unique farsighted stable stable does not satisfy any notion of maximality. But in a negotiation setting, there aren't states as in the example with a highly restricted, tree-like structure describing possible moves. States are combinations of coalition structures and proposed payoff allocations, and 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*. These properties are described as A and B in Section 3.1, and the main result is stated as Theorem 1. The theorem is useful because the identification of farsighted 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 show that 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 observed, 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 it may satisfy strong maximality but not absolute maximality.[3] Coalitional games, however, have more structure and as our positive results show, history dependence allows us to establish *absolute* maximality in a variety of cases. 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

[3]Although they are concerned with history independent processes, these examples also apply to history dependent processes.

time, these properties are sufficient and not necessary for absolute maximality, as shown in Example 3.

In Section 3.5.1 we show that our positive conclusions can be further sharpened if we require only maximality, rather than absolute maximality.[4] In this case, as Theorem 2 shows, Property B can be dispensed with completely. Moreover, in simple games, even Property A need not be checked for any farsighted stable set.

2. Maximal Farsighted Stability

2.1. **Coalitional Games.** A coalitional game, or a characteristic function game, is described by a finite set N of players and a mapping V that assigns to each coalition S (a nonempty subset of N) a closed set of feasible payoff vectors $V(S) \subseteq \mathbb{R}^S$. Normalize the game so that singletons obtain zero, and assume that all coalitions can get nonnegative but bounded payoffs. So $V(S)$ is some nonempty compact subset of $\mathbb{R}^S_+$. A transferable utility (TU) game is one in which each coalition S has a *worth* $\nu(S)$ and $V(S) = \{u \in \mathbb{R}^S \mid \sum_{i \in S} u_i \leq \nu(S)\}$.

2.2. **States and Effectivity.** A *state* is a coalition structure π and a payoff profile u feasible for that structure. A typical state x is therefore a pair (π, 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 resulting 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 $u(x)_T = u(y)_T$.

(E.2) For every state x, coalition S, partition μ 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 $u_T(y) = v$.

Condition E.1 grants coalitional sovereignty to the untouched coalitions: the formation of 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 degree of sovereignty to the moving coalition. It says that if S wants to move from a going state, it can do so by reorganizing itself (breaking up into smaller pieces if it so wishes, captured by the sub-structure μ), 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.

[4] The farsighted stable set described in Example 1 is maximal but not absolutely maximal.

2.3. **Farsighted Stability.** A *chain* is a finite collection of states $\{y^0, y^1, \ldots, y^m\}$ and coalitions $\{S^1, \ldots, S^m\}$, 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 $k = 1, \ldots m$, $u(y)_{S^k} \gg u(y^{k-1})_{S^k}$. The 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:

(i) *Internal Farsighted Stability*. No state in F is farsightedly dominated by another state in F;

(ii) *External Farsighted Stability*. A state not in F is farsightedly dominated by some state in F.

Observe that farsightedness does not impose any optimization on coalitional moves, except for requiring that coalitions must be eventually better off participating in the chain rather than not participating at all. Below, we impose stringent maximality requirements.

2.4. **Absolutely Maximal Farsighted Stable Sets.** To incorporate the notion of maximality in a farsighted stable set, we will "embed" that set into an ambient history-dependent *negotiation process*. To this end, define a *history* h to be a finite sequence of states (where any change of state must be feasible), along with the coalitions that generate any state transitions. If there is no move, the empty coalition is recorded. An *initial history* is just a single state.[5] Let $x(h)$ be the last state in history h. A *negotiation process* is a map σ 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 set $S(h)$ may be empty; i.e., "nothing happens.") In this way, given any history h, σ induces a continuation chain.

A state x is *absorbing* under the process σ if at any history h with $x(h) = x$, $y(h) = x(h) = x$. That is, once at x the continuation chain displays x forever. Say that σ is an *absorbing process* if its continuation chain must terminate in an absorbing state starting from any history.[6] For every absorbing process σ and history h, let $x^\sigma(h)$ denote the absorbing state reached from h. Say that an absorbing process σ is *coalitionally acceptable* if for each history h, if $S(h)$ is nonempty, then $u_S(x^\sigma(h)) \geq u_S(x(h))$. Finally, call an absorbing process σ *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.

[5]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.

[6]That is, there exists k such that $y^{(t)}(h) = x$ for all $t \geq k$, where $y^{(t)}$ is defined recursively in the obvious way.

A farsighted stable set F is *absolutely maximal* if it can be embedded in some absorbing, coalitionally acceptable, and absolutely maximal process σ; that is,

(i) F is the set of all absorbing states of σ.

(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 σ 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 replacement of 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, σ does more: it prescribes a continuation chain for *all* histories, not just the ones described in condition (ii) above. It is necessary to do this, because we need a setting where the counterfactual consequences of alternative actions can be discussed. That requires us to consider deviations from ongoing chains, deviations from deviations, and so on; σ handles all these.

The requirement that σ 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.[7] Absolute maximality is arguably the strongest form of maximality that one could insist on. These distinctions have bite, as illustrated by Examples 1 and 4 below.

Our definition also asks that σ be absorbing: that a negotiation process must ultimately terminate.[8]: 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.[9] It is a joint condition on any starting point and the final

[7]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).

[8]It implies a bit more: it asks for an absorbing state to be absorbing after *every* history leading to it. This property does not follow from rationality per se

[9]Without coalitional acceptability, it is possible to have negotiation processes that return to some single state x from any history, however unpalatable it might be for some or perhaps all players. (No one-shot deviation can be improving.) That is absurd, because some coalition could be better off by refusing to go.

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. This extension is discussed in the Online Appendix.

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.

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 to internal stability. In fact, in the classical literature starting with von Neumann and Morgenstern (1944) and including Harsanyi (1974), it *is* internal stability.[10]

Nevertheless, it is possible to lock two coalitions into a coordination failure so that coalitional acceptability applies to neither of them. Some of these outcomes can be eliminated by perturbing our model so that it applies to negotiations with discounted payoffs in real time, as in Konishi and Ray (2003) and Ray and Vohra (2014). But we find it easiest to impose coalitional acceptability directly on the process.

[10] 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'.

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

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

Section 3.3 proves the theorem. Here, we examine Properties A and B in some applications. Later, in Section 3.4, we examine how tightly these properties are connected to the outcome of interest (the maximality of the stable set). It is to be noted, first and foremost, that Properties A and B are not assumptions, but rather characteristics of an endogenous outcome — the farsighted stable set — to be checked.

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.

The usefulness of Remark 1 depends on the existence of single payoff farsighted stable sets. Ray and Vohra (2015) characterize single-payoff farsighted stable sets based on the notion of a *separable* payoff allocation. 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) show that $[u]$ is a single payoff farsighted stable set if and only if u is separable.

The significance of Remark 1 therefore rests on identifying games that possess separable payoff allocations. As shown in Ray and Vohra (2015), the interior of the core is contained in the set of separable allocations, which are themselves contained in the coalition structure core. Every game in which the interior of the core is nonempty therefore possesses a single payoff farsighted stable set. However, these inclusions can be strict, so the non-emptiness of the core is not generally sufficient for the existence of separable allocations.[11] Known special cases in which a 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 now turn to an important case — exchange economies— in which all core allocations may not be separable but there is a distinguished one that is.

Among the most fruitful economic applications of coalitional games have been those relating to exchange economies. It is therefore of some significance that we can provide

[11] This is the case, for instance, in Example 4.

reasonable, sufficient conditions for a competitive equilibrium to yield a separable payoff allocation. Assumptions that ensure the existence of a competitive equilibrium are, of course well known; see, for example, Mas-Colell, Whinston and Green (1995).

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, $u_i : X_i \to \mathbb{R}$ 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 ξ_i denotes i's commodity bundle, and $p \in \mathbb{R}^l_+$ is 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) $\sum_{i\in N} \xi_i = \sum_{i\in N} \omega_i$.

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 ξ_i such that $u(\xi_i') > u_i(\xi_i)$; and (b) strictly convex: $u_i(\xi_i') \geq u_i(\xi_i)$ and $\xi_i' \neq \xi_i$ implies that $u_i(t\xi_i' + (1-t)\xi_i) > u_i(\xi_i)$ for all $t \in (0, 1)$.

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. With locally non-satiated and strictly convex preferences, the payoff profile u corresponding to any competitive equilibrium of an exchange economy is separable. By Ray and Vohra (2015, Theorem 2), $[u]$ is a single-payoff farsighted stable set, and Remark 1 applies.[12]

To prove this Remark, consider a competitive equilibrium $(\{\xi_i\}, p)$. We will show that $u \equiv \{u_i(\xi_i)\}$ is separable. Suppose there is a coalition S such that $u_S \in V(S)$. This means that there exists a feasible allocation ξ' for the economy with agent set S such that $\sum_{i\in S} \xi_i' = \sum_{i\in S} \omega_i$ and $u_i(\xi_i') \geq u_i$ for all $i \in S$. Since preferences are locally non-satiated, condition (i) of a competitive equilibrium implies that $p \cdot \xi_i' \geq p \cdot \omega_i$ for all $i \in S$. In fact, it must be the case that

$$p \cdot \xi_i' = p \cdot \omega_i \text{ for all } i \in S, \tag{1}$$

otherwise we contradict the feasibility condition $\sum_{i\in S} \xi_i' = \sum_{i\in S} \omega_i$. Next, we claim that $\xi_i' = \xi_i$ for all $i \in S$. If not, there is some $i \in S$ with $\xi_i' \neq \xi_i$. By the strict convexity

[12]Greenberg, Luo, Oladi and Shitovitz (2002) study what they call "the sophisticated stable set" of an exchange economy. This is based on a version of the Harsanyi stable set in which every step of a blocking chain is also required to be a myopic objection. The core is consequently a subset of the sophisticated stable set. In general, therefore, the competitive equilibrium is not a single-payoff sophisticated stable set. A second difference between the farsighted stable set and the sophisticated stable set is that the latter, as in Harsanyi (1974), allows a deviating coalition to choose *any* feasible payoff for the complementary coalition, a notion that is critiqued and dropped in Ray and Vohra's (2015) development of the farsighted stable set. Last but not least, our focus here is on maximality.

of u_i, there is a strict convex combination of ξ_i' and ξ_i which is strictly preferred to ξ_i. By (1) it is also affordable. But this contradicts condition (i) of a competitive equilibrium. It follows that $\sum_{i\in S}\xi_i = \sum_{i\in S}\xi_i' = \sum_{i\in S}\omega_i$. Because $\sum_{i\in N}\xi_i = \sum_{i\in N}\omega_i$, this implies that $\sum_{j\in N-S}\xi_j = \sum_{j\in N-S}\omega_j$, and $u_{N-S}\in V(N-S)$; i.e., u is separable.

Although Remark 2 applies even to economies in which the interior of the core is empty, it does depend crucially on preferences being *strictly* convex. In fact, the characteristic function game in Example 4 below can be derived from a three–consumer exchange economy with prefect complements (and preferences that are convex but not strictly convex). In that economy, the competitive allocation — the only one in the core — is *not* separable.

We now move on to a consideration of farsighted stable sets with nonsingleton payoffs. We begin with a straightforward observation:

REMARK 3. If every state in a farsighted stable set has the grand coalition as the associated coalition structure, then Property B is satisfied.

That is because Property B is then equivalent to myopic internal stability, which is implied by farsighted internal stability.

The general structure of farsightedness is yet to be fully understood, but we know quite a bit for particular classes of games. Specifically, Ray and Vohra (2015) provide a full analysis of *simple games*, which are TU games in which each coalition S is either "winning" ($\nu(S)=1$) or "losing" ($\nu(S)=0$), and if a coalition is winning, then its complement is losing. Despite the simple-sounding nomenclature, simple games describe a rich class of situations: parliaments, bargaining institutions, and committees have been studied with this device.[13] 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 $i\notin W(x)$. The state with no winning coalition is referred to as the *zero state*.

In all such situations Property B is redundant:

REMARK 4. Every farsighted stable set in a simple game satisfies Property B.

This remark is a consequence of farsighted internal stability. To see it, suppose there is a farsighted stable set F for which Property B fails. Then there are states a and b in F and a coalition T such that $u_T(b)\gg u_T(a)$ and $\sum_{i\in T}u_i(b)\le\nu(T)$. This implies that T is a winning coalition, but then the complement of T is losing. So at state a, T

[13]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.

can precipitate the zero state (by breaking up into singletons), counting on the winning coalition for state b to move to b, making T better off. Therefore b farsightedly dominates a, which contradicts the farsighted internal stability of F. So Property B must hold.

To understand whether Property A holds for simple games, it is useful to distinguish between two subclasses. Define a *veto player* as an individual with a losing complement (she can single-handedly precipitate the zero state). If the set of all veto players is winning, say that the game is oligarchic. Oligarchic games have singleton farsighted stable sets (Ray and Vohra 2015, Theorem 3), which trivially satisfy Property A. Otherwise, the game is non-oligarchic, and now there are no singleton farsighted stable sets.

And yet, Property A is satisfied by a class of sets that played a central role in von Neumann and Morgenstern's analysis of stability: *discriminatory sets*, to use their terminology. These are sets of the form

$$D(K, c) = \{x \in X \mid u_i(x) = c_i \text{ for } i \in K\}$$

for some fixed player set $K \subseteq N$ and associated payoff vector $c \in \mathbb{R}^K$. Those in K, the "discriminated players," each get a fixed amount, while the remaining surplus is divided arbitrarily among the remainder, the "bargaining players."

REMARK 5. Every discriminatory farsighted stable set satisfies Property A. As shown in Ray and Vohra (2015, Theorem 5), such sets exist in every non-oligarchic simple game that has a minimal veto coalition with no indispensable members.[14]

To see why this is true, let $a, b \in D(K, c)$, with $u_j(b) > u_j(a)$ for some j. Clearly, $j \notin K$, which means that there is $z \in D(K, c)$ with $u_k(z) = c_k$ for all $k \in K$, $u_i(z) \geq u_i(b)$ for all $i \neq j$, and $u_j(z) = 0$. Therefore $D(K, c)$ satisfies Property A.

For simple games, another set — with a discrete collection of payoffs — is a potential candidate for a stable set (von Neumann and Morgenstern 1944). For any vector $m \in \mathbb{R}^N$ with $m \gg 0$ and $\sum_{i \in S} m_i = 1$ for every minimal winning coalition S, define

$$Z(m) = \{x \in X \mid S(x) \text{ is minimal winning and } u_i(x) = m_i \text{ for } i \in S\}$$

to be a *main simple set*. von Neumann and Morgenstern (1944) showed that if a game is *strong* — every coalition is either winning, or its complement is — then the set of utility profiles corresponding to a main simple set is a vNM stable set. Ray and Vohra (2015) showed that a main simple set (of a strong, simple game) is a farsighted stable set.

In general, a main simple set may not satisfy Property A, as we will see in Example 1. But an important subclass of simple games yields a different answer. Say that a simple game is *symmetric* if there is some k, where $(n+1)/2 \leq k \leq n$, such that every coalition with k players is a minimal winning coalition. (Supermajority games have this property.) Every symmetric simple game has a main simple set $Z(m)$, with $m_i = 1/k$

[14]That is, any coalition member can be replaced by any outsider, and the coalition would remain veto.

for all i. Observe that a symmetric game may not be strong. Yet its main simple set is indeed a farsighted stable set, though it may fail to be a vNM stable set.[15]

Moreover, if the game is non-oligarchic ($k < n$), $Z(m)$ satisfies Property A. To see this, suppose a and b are in $Z(m)$, with $u_i(b) > u_i(a)$. This implies that there is a minimal winning coalition S such that $i \in S$, $u_i(b) = 1/k$, and $u_j(a) = 0$ for all $j \notin S$. Since $k < n$, there exists $j \notin S$. Let $S' = S - i + j$. Given the symmetry of the game, this is a minimal winning coalition and the corresponding state in $Z(m)$, say z, has the property that $u_i(z) = 0$ and $u_j(z) \geq u_j(b)$ for all $j \neq i$, which yields Property A. To summarize:

REMARK 6. Any non-oligarchic, symmetric simple game possesses a main simple set which is farsighted stable and satisfies Property A.

Collecting Remarks 1–6, we have:

COROLLARY **1.** *A farsighted stable set F is absolutely maximal in any of the following circumstances:*

(i) *F is a single-payoff farsighted stable set; e.g., a competitive equilibrium of an exchange economy with strictly convex preferences,*

(ii) *F satisfies Property A and for every $x \in F$, $\pi(x) = N$,*

(iii) *F satisfies Property A and the game is simple,*

(iv) *F is a discriminatory set of a simple game,*

(v) *F is a main simple set of a symmetric, simple game.*

(*Sufficient conditions for the existence of farsighted stable sets of the form stated in* (i), (iv) *and* (v) *have been noted above.*)

3.3. **Proof of Theorem 1.** 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

[15]Consider the non-strong, symmetric simple game with $n = 5$ and $k = 4$. Then $Z(m)$ is not vNM stable: the state x with $u(x) = (1/3, 1/3, 1/3, 0, 0)$ has no objection from $Z(m)$. However, there is a farsighted objection through the zero state initiated by players 4 and 5 (a veto coalition), leading to $Z(m)$. Indeed, farsighted stability holds for all such games. To see why $Z(m)$ satisfies farsighted external stability, consider $x \notin Z(m)$. Clearly, $S = \{i \in N \mid u_i(x) \geq 1/k\}$ must then be a losing coalition. If the complement of S, $N - S = \{i \in N \mid u_i(x) < 1/k\}$, is winning, any minimal winning coalition in $N - S$ can (myopically) block x with a state in Z_m. Otherwise, because S is losing, $N - S$ is a veto coalition and can farsightedly block x by first precipitating the zero state and then moving (via a suitable minimal winning coalition) to obtain $1/k$ for all its members. The main simple set also satisfies farsighted internal stability (under a mild monotonicity restriction on the effectivity correspondence); see Ray and Vohra (2015) or Dutta and Vohra (2017). Thus $Z(m)$ is a farsighted stable set. These arguments can be extended to show that a main simple set of any (not necessarily symmetric) simple game is a farsighted stable set, although absolute maximality cannot be assured, as shown by Example 1.

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.

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

(i) *$y^0 = x$ and $y^m = y$; and*

(ii) *S^i and S^j are disjoint for all i and j between 1 and $m-1$.*

Proof. Set $y^0 = x$ and $S^1 = \tilde{S}^1$ and recursively, let $S^k = \tilde{S}^k - \cup_{t<k}\tilde{S}^t$ for all $k = 2, \ldots, \tilde{m}-1$. When coalition S^k moves, it does so by breaking into singletons. So, for any $k = 1, \ldots, \tilde{m}-1$, the corresponding coalition structure, π^k, is such that all players in $\cup_{t<k}\tilde{S}^t$ are in singletons, and (by Condition E.1) 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 $\pi^k, k = 1, \ldots, \tilde{m}$ in the new chain, it follows from Conditions E.1 and E.2 that

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

Let the associated states be $y^k = (u^k, \pi^k)$ for all $k = 1, \ldots, \tilde{m}-1$, and $y^{\tilde{m}} = y$. It is possible that for some stages $k < m$, S^k as defined is empty and the succeeding state y^{k+1} is identical to y^k.[16] 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.

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.

[16] This will happen when a new coalition belongs to the union of previous coalitions in the chain.

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, $\mathbf{c} = \{y, y^1, \ldots, y^{m-1}, y^m\}$, $\{S^1, \ldots, S^m\}$, where $(y^m, S^m) = b$. 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.

We now consider two cases:

Case 1. Some subset W of T moves *only* in the final step from y^{m-1} to b, and so is part of the coalition S^m. Pick any $j \in W$. Modify the original blocking chain by adding an extra step at y^{m-1} in which $W - j$ breaks up into singletons and moves from y^{m-1} to y'. Note that $\pi(y')$ has all the same coalitions as $\pi(y^{m-1})$, except that W appears as singletons, and $u_i(y') = u_i(y^{m-1})$ if $i \notin W$, while $u_i(y') = 0$ if $i \in W$. With y' as an added step between y^{m-1} and y^m we have a new chain $\mathbf{c}' = \{y, y^1, \ldots, y^{m-1}, y', y^m\}$, $\{S^1, \ldots, S^{m-1}, W - j, S^m\}$, with $(y^m, S^m) = b$.[17] Clearly, this new chain is also a blocking chain. The critical feature of this new blocking chain is that at state y' player j has yet to move and $u_j(y') = 0$.

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 $\mathbf{c}'$ by replacing the terminal state with z to construct the chain $\bar{\mathbf{c}} = \{y, y^1, \ldots, y^{m-1}, y', z\}$, $\{S^1, \ldots, S^{m-1}, W - j, N\}$. Since $u_i(z) \geq u_i(b)$ for all $j \neq i$, this chain clearly satisfies the acceptability conditions for all players in $N - j$. Since player j only moves at the last step, from y' to z, and $u_j(y') = 0$, the acceptability condition also holds for player j. Thus, $\bar{\mathbf{c}}$ is a coalitionally acceptable chain from y to z such $u_j(z) \leq u_j(a)$.[18]

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

[17]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 $u_j(y') = 0$.

[18]Any player other player j who was an active mover in the original chain $\mathbf{c}$ (i.e. any player in $S^m - j$) will also gain *strictly* in the chain $\bar{\mathbf{c}}$. There are precisely two reasons why $\bar{\mathbf{c}}$ may not be a blocking chain: (i) some players in $N - S^m$, who were not involved in making a move in $\mathbf{c}$, may be assigned to a different coalition in $\pi(z)$ and may not gain (strictly) in following the path to z, (ii) it's possible that $u_j(y') = u_j(z) = 0$ so j does not experience a strict improvement.

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.[19] 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 σ 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 σ, 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 σ 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 σ 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 σ 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 $c(h_K)$: 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

[19] Formally, replace S^k by $S^k - j$.

chain $\mathbf{c}'$ and assign it to the history h, defining σ accordingly at h. This last step ensures that for no h can a coalition profitably deviate from the path prescribed by $\mathbf{c}(h)$.

Proceeding recursively in this way, we define $c(h)$ for every h, along with the accompanying $\sigma(h)$. Clearly, σ embeds F and is coalitionally acceptable.[20]. 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 Importance of Properties A and B.** As discussed, there is a sizable class of games with farsighted stable sets satisfying Properties A and B. In this section, we ask: (i) Are these properties nontrivial restrictions, or are they always satisfied by any farsighted stable set? (ii) Even if they are not, might Theorem 1 still be valid without them? (iii) Are these properties necessary for a farsighted stable set to be absolutely maximal? The answer to each of these questions 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 $\{3, 4\}$.[21] This game is strong; we set up a main simple set for it. To this end, 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 4, 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. Indeed, as we shall now show, F is generally not absolutely maximal,[22] which also indicates that Property A cannot be freely removed from the statement of Theorem 1.

To make this point, we impose a "monotonicity condition" on the effectivity correspondence E.[23] Assume that if a winning coalition loses some members but remains winning, the resulting nonnegative surplus (captured from the departing members) is shared equally among the players that remain. Now suppose by way of contradiction that there is an absorbing σ that embeds F and satisfies coalitional acceptability

[20] In fact, barring case (iii), every h with $x(h) \notin F$ is assigned to a *blocking* chain terminating in F.

[21] This game can also be represented as a weighted majority game where the players' weights are $(1, 1, 2, 3)$ and a coalition is winning if and only if its aggregate weight is more than 3.5.

[22] We do know from Dutta and Vohra (2017) that F is strongly maximal.

[23] See Ray and Vohra (2015) for a more general version that applies to all games.

and absolute maximality. Consider state x with $u(x) = (0, 0, 0.36, 0.64)$ and winning coalition $W(x) = N$. Because $x \notin F$, there is $x' \in F$ that farsightedly dominates it; i.e., σ leads from history $h = \{x\}$ to x'. Ray and Vohra (2015, Lemma 1) show that there are just two possibilities: either (i) x' myopically dominates x, or (ii) $W^+ = \{i \in N \mid u_i(x') > u_i(x)\}$ and $W(x) - W^+$ are both losing coalitions.[24] But $W(x)$ equals N and our game is strong, so the second option must be eliminated here. It follows that (i) is true: x' myopically dominates x. But the only two states in F that do so are $x' = ((1/3, 0, 0, 2/3), \{1, 4\})$ or $x' = ((0, 1/3, 0, 2/3), \{2, 4\})$. In either case, $u_3(x') = 0$. We use this last fact to argue that player 3 can profitably deviate from the stipulated move at x (to x'), thus violating absolute maximality.

Suppose player 3 leaves the grand coalition at x resulting in state y. Note that the residual coalition, $\{1, 2, 4\}$ is winning. Given that the residual players share equally in the surplus released by 3's departure, $u(y) = (0.12, 0.12, 0, 0.76)$. Since $y \notin F$, σ must prescribe a continuation that is coalitionally acceptable. Using the same kind of argument as in the previous paragraph, it can be shown that $x^\sigma(y) = ((1/3, 1/3, 1/3, 0), \{1, 2, 3\})$.[25] Player 3 can therefore gain by interfering in this way with any process that attempts to proceed from x to x'. In other words, F does not satisfy absolute maximality.

Interestingly, there is a non-elitist veto coalition in Example 1, namely $\{1, 2, 3\}$.[26] By Ray and Vohra (2015, Theorem 5), there exists a discriminatory farsighted stable set, for example $D(\{4\}, 0.1)$, in which player 4 receives 0.1 and players 1, 2 and 3 receive any arbitrary division of 0.9. By Remark 5 in Section 3.2, this set is absolutely maximal. Example 1 therefore illustrates how absolute maximality can refine the set of farsighted stable sets for 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 $\nu(S)$. (Such a game is referred to as a hedonic game). A few words of explanation will help. Players 1 and 5 are symmetric, so are players 2 and 4. These are the four players whose positive payoffs vary across different states. 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 herself; player 6 is completely indifferent throughout.

---

[24] In case (ii), W^+ can precipitate the zero state by leaving $W(x)$, followed by a move by $W(x')$ to x'.

[25] For example, if $x^\sigma(y) = ((1/3, 0, 0, 2/3), \{1, 4\})$ coalitional rationality implies that in the first step player 1 must leave $W(y)$, resulting in the state $y' = ((0, 0.18, 0, 0.82), \{2, 4\})$. But from y' it is not possible, by coalitional rationality, to end up at $((1/3, 0, 0, 2/3), \{1, 4\})$.

[26] As a minimal winning coalition of a strong game it is of course a minimal veto coalition. It is non-elitist because it remains a veto coalition if we replace any of its members with player 4.

Formally, the coalitional game is described as follows:

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

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 π_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.[27]

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.

		Payoffs to Players					
States	Structures	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.

We first show that F is a farsighted stable set. Figure 1 shows all the payoff equivalent states, with arrows indicating the states in F that farsightedly dominate a state not in F.

To see that F satisfies external farsighted stability: A state with payoff $(3,3,0,3,3,0)$ is dominated by one in X^3 through coalition $\{2,3,4\}$. The state with payoff $(3,3,0,0,0,0)$ is directly dominated by one in X^1 through $\{3,5\}$ and by one in X^3 through $\{2,3,4\}$. A state in X^8, with payoff profile $(1,0,2,0,1,0)$, is farsightedly dominated by a state in X^1 through the formation of coalition $\{1,2\}$, and also by a state in X^2 through coalition $\{4,5\}$. It is easy to see from Figure 1 that other states not in F are also farsightedly dominated by some state(s) in F.

27 For instance, $\Pi^0(\{1,2,3\}) = \{(\{1\},\{2\},\{3\}),(\{1,2,3\}),(\{1\},\{2,3\})\}$.

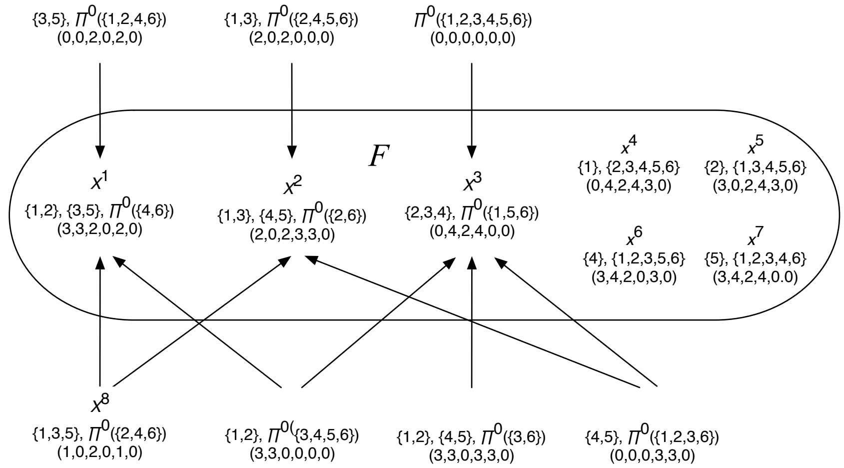

FIGURE 1. External Farsighted Stability of F in Example 2

To see that F satisfies internal farsighted stability: First observe that states x^4, x^5, x^6 and x^7 cannot dominate any other state (these states are in F only because they cannot be dominated by a state in $X^1 \cup X^2 \cup X^3$). This is so because such a state can emerge in only one of two ways: either a singleton precipitates it by leaving the grand coalition or it involves the active participation of player 6. Either case is inconsistent with farsighted dominance because both the singleton as well as player 6 receive 0. Secondly, none of these states can be farsightedly dominated by any other state. All players except for the excluded singleton are receiving the maximum possible payoff. Only the singleton has an incentive to change the state, but on her own she is powerless to do so. Thus, in checking internal stability we only need to compare states in X^1, X^2 and X^3.

From X^1 the only players who could gain by ending up at a state in X^2 are players 4 and 5. They can't move there directly. They could form a coalition of their own, resulting in payoffs $(3,3,0,3,3,0)$, but that can only be dominated by a state in X^3, not one in X^2, resulting in a payoff of 0 to player 5, which is of course not a farsighted improvement for $\{4,5\}$. Player 5 could exit coalition $\{3,5\}$ resulting in payoffs $(3,3,0,0,0,0)$ but from there the only possible moves are into X^1 or X^3, again making it impossible for player 5 to gain.

A state in X^1 cannot be farsightedly dominated by one in X^3 because any such move must begin by player 2 leaving coalition $\{1,2\}$ which results in payoffs $(0,0,2,0,2,0)$ from which the only further move that is possible is to X^1 or to X^2, not X^3, because players 3 or 5 the only ones who could initiate a move to X^3 have no interest in doing so. A similar argument shows that no state in X^2 can be farsightedly dominated by another state in F. Finally, note that at a state in X^3, all the non-zero payoff players are getting the highest possible amount and they together belong to one coalition, so no profitable deviation is possible.

This completes the proof that F satisfies farsighted internal stability.

Next, we show 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).[28] 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.

Finally, we can show that F is not absolutely maximal; that is, it cannot be embedded in an absorbing, coalitionally acceptable, absolutely maximal process. This argument relies crucially on the fact that from a state in X^8 the *only* possible farsighted blocking chain runs to states in X^1 or X^2, not to a state in X^3. This is so because the only players who would prefer to have a state in X^8 replaced by one in X^3 are 2 and 4, but without the active participation of player 3 they are unable to carry out such a move.

Now consider any process that satisfies embedding of blocking chains and absorption into F. Take the history consisting of a single state in X^8. Since the only blocking chains from such a state are into X^1 or X^2, the continuation must be a single step into X^1 (through coalition $\{1, 2\}$) or into X^2 (through coalition $\{4, 5\}$). In the former case, coalition $\{4, 5\}$ has a profitable deviation into X^2 and in the latter coalition $\{1, 2\}$ has a profitable deviation into X^1. Thus, F is not absolutely maximal, which also shows that Property B cannot be dispensed with in our main theorem.

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. See Section 3.5.2 for a different remark on the folk theorem.

While Examples 1 and 2 show that neither Property A nor Property B can be dropped from the statement of our main Theorem, these properties are not necessary for a farsighted stable set to be absolutely stable. This is shown through the following example.

[28] 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.

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

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

It can be shown that $F = \{x^1, x^2\} \cup X^3$, described in Table 2, is 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.

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 $u_5(x) \geq 2$.

The farsighted stability of F follows from arguments we already provided in the discussion of Example 2. In Figure 2 the arrows from states outside F represent a process that embeds F. We leave it to the reader to check that it satisfies absolute maxlimality.

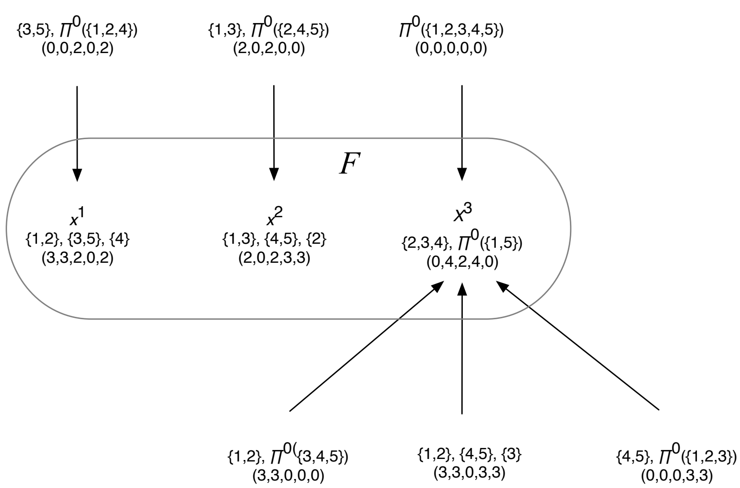

FIGURE 2. Absolute Farsighted Stability of F in Example 3

3.5. **Other Approaches to Maximal Farsightedness.** We seek conditions under which a farsighted stable set might satisfy maximality. The underlying idea is to begin with a solution concept that is the natural farsighted extension to a classical notion — vNM stability — and attempt to embed that concept within an ambient negotiation process satisfying the desideratum of absolute maximality. In this section, we discuss two alternatives. (A third alternative, which explores the use of blocking chains following *all* histories, is explored in the Online Appendix.)

3.5.1. *Weakening Absolute Maximality.* A farsighted stable set is *maximal* if it can be embedded in an absorbing, coalitionally acceptable process σ satisfying conditions (i) and (ii) in the main text, and:

[Maximality] At no history h does there exist a state y with $S(h) \in E(x(h), y)$, such that $u_{S(h)}(x^\sigma(h, y, S(h))) \gg u_{S(h)}(x^\sigma(h))$.

The difference from absolute maximality is that maximality requires only the moving coalition at any history to not have a better move. Under absolute maximality *no* coalition can have a better move. Our weaker notion can be thought of as one that respects protocols: interventions by other coalitions are ignored. If we entertain this weakening, Property B can be dropped from Theorem 1. Moreover, in simple games even Property A can be dropped:

THEOREM **2.** *If a farsighted stable set satisfies Property A then it is maximal.*

The proof relies on a version of Lemma 2 that does not invoke Property B, and is provided in the Online Appendix.

In simple games even Property A can be dropped from Lemma 5. Fix states x, y, a, b and a coalition T as in the statement of Lemma 5 and suppose that $u_T(b) \gg u_T(a)$. In a simple game T must be a losing coalition; if it is a winning coalition, it could have gone from a to b on its own, contradicting internal stability. But a losing coalition that was supposed to move from x to y' cannot move anywhere other than y' (because of coalitional sovereignty), so in fact it has no available "deviation", and the conclusion of Lemma 5 follows. As in the proof of Proposition 1 (see Online Appendix), we can embed a farsighted stable set in a process satisfying maximality.

REMARK 7. Every farsighted stable set of a simple game can be embedded in a process satisfying maximality.

The Online Appendix proves a stronger version of this as part of Proposition 2.

3.5.2. *The Absorbing States of a Negotiation Process.* Alternatively, one might directly seek to understand the absorbing states of a negotiation process, without asking that any existing solution concept be embedded in it. Such an approach is followed in Konishi and Ray (2003) and Dutta and Vohra (2017), along with the additional restriction that the negotiation process σ is Markovian or history independent: for any two histories h and h', $x(h) = x(h')$ implies that $\sigma(h) = \sigma(h')$.[29] A comparison of these two approaches is instructive.[30] We turn to an example that makes the following points:

1. While unclear from its original definition (Harsanyi 1974, Ray and Vohra 2015), a farsighted stable set is fundamentally a history-dependent object. There is little hope of being able to embed a farsighted stable set in a Markovian process, and this is so even if we ignore the maximality requirement. On the other hand, as our main results demonstrate, permitting history dependence can make it possible to embed a farsighted stable set in a process that is *absolutely maximal*.

2. Absolute maximality can be a more stringent requirement than maximality or strong maximality even if the focus is on absorbing states that are not necessarily a farsighted stable set.[31]

EXAMPLE 4. A three-player veto game: $N = \{1, 2, 3\}$, $\nu(N) = \nu\{1, 2\} = \nu\{1, 3\} = 1$ and $\nu(S) = 0$ for all other S.

Ray and Vohra (2015) show that *every* farsighted stable set of this game is a discriminatory set of the form $Z_a = D(\{1\}, a)$ in which player 1 receives a fixed payoff $a \in (0, 1)$ and the remaining surplus is divided in any arbitrary way between players 2 and 3. (In fact, for every $a \in (0, 1)$, Z_a is a farsighted stable set.) It follows from Theorem 1 and Proposition 2 (ii) that any such set can be embedded in an absolutely maximal process.

But this result depends crucially on allowing the process to be history dependent. As Dutta and Vohra (2017) point out, a set of this form cannot be supported by a Markovian process that is consistent with farsighted external stability. The farsighted external stability of Z_a implies that from any state x with $u(x) \gg 0$ and $u_1(x) > a$ there is a blocking chain ending in Z_a. It can be shown that any such chain involves players 2 and 3 leaving the grand coalition at state x, resulting in the zero state. This is then followed by a move by N to a state in Z_a; see Ray and Vohra (2015) for details. It turns out that the last step of any such blocking chain *must* depend on the history.

To see this, suppose σ is a Markovian process that defines, for every state not in Z_a, a blocking chain that ends in Z_a. Consider the zero state, x^0, and suppose σ prescribes a path from x^0 that ends at $y \in Z_a$. Since the process is Markovian, this is the continuation paths for *all* histories where the current state is x^0. Consider x such that $u(x) \gg 0$ and $u_1(x) > a$. As already observed, any blocking chain from x leading into Z_a must involve

[29] History-dependent versions of these solutions are studied in Hyndman and Ray (2007), Ray and Vohra (2014) and Dutta and Vartiainen (2018).

[30] We are thankful to an anonymous referee for urging us to do this.

[31] With respect to farsighted stable sets this point has already made; through Example 1 for simple games and through the examples in Dutta and Vohra (2017) for abstract games.

coalition $\{2,3\}$ moving to x^0, followed by a move by N to y (with $u(y') \gg 0$). Since $u(y') \gg 0$, we can find x such that $u_1(x) > a$, $u_2(x) > u_2(y)$ and $u_3(x) > 0$. The process must specify a blocking chain from x to a state in Z_a. But any such blocking chain must be one in which $\{2,3\}$ first moves to x^0 followed by a move to y. Since $u_2(x) > u_2(y)$ player 2 cannot gain. In other words, the path prescribed by σ from x is not a blocking chain, a contradiction.

This example also illustrates the difference between our approach and one that directly examines the absorbing states of a process, without seeking to embed a particular solution. Consider the Dutta and Vohra (2017) notion of a *strong rational expectations farsighted stable set* (SREFS) which is defined to be the set of absorbing states, Z, of a Markovian process σ that satisfies strong maximality, as well as both internal and external stability when blocking chains are restricted to be consistent with σ. In particular, if a coalition moves from a state in Z, it cannot eventually gain *provided* the continuation following this move is given by σ. There may, however, exist a farsighted blocking chain that is not consistent with σ, and for this reason Z may not be a farsighted stable set.[32] Indeed this is a feature of the present example. Dutta and Vohra (2017) show that there is a SREFS consisting of states with payoffs $(\{a+b,b,0),(a+b,0,b),(a,b,b)\})$, where $a \in (0,1)$ and $b = (1-a)/2$. Of course, this is not a discriminatory set in which player 1 gets a fixed payoff, so it cannot be a farsighted stable set.[33]

While this SREFS satisfies strong maximality, it does not satisfy absolute maximality. To see this, consider the state x with $\pi(x) = N$ and $u(x) = (a+b-1/3\epsilon, b-2/3\epsilon, \epsilon)$. Coalition $\{1,2\}$ can block this is one step to get payoffs $(a+b,b)$. No coalition that includes either player 1 or 2 can construct a better deviation, as is required for strong maximality. However, absolute maximality may not hold because of a deviation by player 3. Suppose that a departure by player 3 results in the other two sharing the extra surplus equally. Now, if player 3 leaves the grand coalition, the new state leaves player 2 with a payoff strictly less than b. This only lead to the zero state followed by N moving to the stationary state with payoffs (a,b,b). Thus, player 3, has a profitable deviation at state x, and the process is not absolutely maximal.

We make a final comment on folk-theorem-like arguments. In Section 3.4, we remarked that there are tight restrictions on the structure of absolutely maximal farsighted stable sets, so it isn't the case that anything goes. That argument carries over to the set of states that comprise any farsighted stable set: "anything doesn't go" because the internal stability of a (farsighted) stable set precludes it from being too inclusionary. However, what would happen under the alternative approach of this section, where we do not insist in embedding any farsighted stable set? Might that span the entire set of feasible

[32] The same is true of the solutions constructed by Dutta and Vartiainen (2018), allowing for history dependence, and using a weaker notion of absorption that we have defined above. Indeed, their solutions may not even satisfy myopic internal stability.

[33] There is a farsighted objection from $((a,b,b),N)$, led by player 1, to $((a+b,b,0),\{1,2\})$.

payoffs? In general, in our model, the answer is still no. That follows from absolute maximality and our notion of an absorbing state which requires once such a state is reached, regardless of the history, it does not change. Together, these two properties imply that a non-core state and a state that (myopically) dominates it can't both be absorbing states.[34] Thus, in general, the absorbing states cannot span the entire set of feasible payoffs.

A loosening of these restrictions could lead to outcomes in which the entire set of payoffs is supportable. Dutta and Vartiainen (2018) consider a weaker notion of absorption or stationarity. A stable outcome in their sense may be stationary for some histories but not for others. The set of such states need not satisfy internal stability even if the process is maximal, and the set of "stable" outcomes can be large. Indeed, they find that in a strictly superadditive game, the set of all strictly positive, feasible payoffs is a "farsighted stable set" in their sense.

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.

[34]In other words, the set of absorbing states must satisfy myopic internal stability.

Dutta, Bhaskar and Hannu Vartiainen (2018), "Coalition Formation and History Dependence", mimeo.

Greenberg, Joseph (1990), *The Theory of Social Situations: An Alternative Game-Theoretic Approach*, Cambridge, UK: Cambridge University Press.

Greenberg, Joseph, Luo, Xiao, Oladi, Reza, and Benyamin Shitovitz (2002), "(Sophisticated) Stable Sets in Exchange Economies," *Games and Economic Behavior* **39**, 54–70.

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.

Mas-Collell, Andreu, Michael Whinston, and Jerry Green (1995), *Microeconomic Theory*, Oxford University Press.

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.

——— (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 ⓡ and 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.

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.

ONLINE APPENDIX [NOT FOR PUBLICATION]

APPENDIX A.1. STRENGTHENING COALITIONAL ACCEPTABILITY

As we argued in Section 3, we restrict off-path histories to satisfy coalitional acceptability, which is a weaker requirement than the strict improvements required by a blocking chain. What if we were to strengthen coalitional acceptability to the requirement that the process define a blocking chain from *every* history ending at a state outside the farsighted stable set? Treating off-path and on-path histories in the same way does have the benefit of simplifying the definition of absolute maximality.

A farsighted stable set F is said to be *strictly embedded* in an absorbing process σ if

(i) F is the set of all absorbing states of σ.

(ii)′ At any initial history h with $x(h) \notin F$, the continuation chain from h is a blocking chain terminating in F.

Condition (ii)′ subsumes our earlier condition (ii) of embedding. It also ensures that σ automatically satisfies coalitional acceptability. We now explore the possibility of extending our result to show that a farsighted stable set can be strictly embedded in a process satisfying absolute maximality.

To strengthen Theorem 1 in this direction, it should be clear from its proof that it suffices 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 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. Indeed, this must be a feature of any state in the farsighted stable that dominates some other state. In a general coalitional game, a state x is said to be *regular* if $u_i(x) > 0$ for every i such that $i \in S \in \pi(x)$ and $u_S(x) > 0$. In a simple game this reduces to the condition that $u_i(x) > 0$ for all $i \in W(x)$, as in Ray and Vohra (2015). That suggests that Property A must be modified 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 $i \neq j$.*

Clearly, every discriminatory farsighted stable set satisfies Property A′.

Modulo the replacement of Property A with A′, our next result shows that Corollary 2 can be strengthened to require strict embedding.

PROPOSITION **1.** *If a farsighted stable set satisfies Property A′, then it can be strictly embedded in a process satisfying absolute maximality in any of the following circumstances:*

(i) F is a single-payoff farsighted stable set (in which case even Property A′ is redundant),

(ii) If the game is a simple game, or

(iii) For every $y \notin F$ there is a blocking chain from y to $x \in F$ with $\pi(x) = N$.[35]

Proof. (i) By Theorem 1 in Dutta and Vohra (2017) a single-payoff farsighted stable set can be strictly embedded in a history-independent (stationary) process satisfying maximality. Clearly *no* coalition can find a profitable deviation when the final payoff is unique, so in this case absolute maximality is also satisfied.

(ii) Suppose F is a farsighted stable set of a simple game. For every state $x \notin F$, fix any blocking chain, $\mathbf{c}(x)$, with $\Psi(x) \in F$ as the terminal state. We can now modify Lemma 2 as follows:

LEMMA **3.** *Consider a farsighted stable set F of a simple game satisfying Property A$'$. 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 $j \in T$.*

Proof. Fix states x, y, a, b and a coalition T as in the statement of the lemma. Note that as the terminal states of a blocking chain both a and b are regular states. If $u_j(b) \leq u_j(a)$ for some $j \in T$ there is nothing to prove, so assume that $u_T(b) \gg u_T(a)$. This means that T is not a veto coalition. Otherwise, we could find a farsighted chain from a to b (with T first moving to the zero state followed by $W(b)$ moving to b), contradicting the farsighted internal stability of F. Since T is not a veto coalition, and therefore not a winning coalition, $u_T(y) = 0$, which clearly means that $y \neq b$ and $y \notin F$. Pick any $j \in T$. By Property A$'$, 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 $i \neq j$. To complete the proof we will now construct a farsighted chain from y to z.

Since T is losing, $u_T(y) = 0$ and $T \cap W(y) = \emptyset$. Given that there is a blocking chain from y to b, by Lemma 1 of Ray and Vohra (2015) there are two possibilities: (i) $W(b)$ moves directly from y to b or (ii) some $S \subset W(y)$ first moves from y to the zero state, followed by a move by $W(b)$ to b. In case (i) consider a move by $W(z)$ to the zero state (by breaking up into singletons) followed by a move by $W(z)$ to z. In case (ii) consider S moving to the zero state followed by $W(z)$ moving to z. All players except possibly j who gained by having y replaced with b, will also gain if y is replaced with z. If player $j \notin W(z)$ we have clearly found a blocking chain from y to z. If $j \in W(z)$, since z is regular, $u_j(z) > 0$ while $u_j(y) = 0$, so player j also gains in moving from y to z, and again we shown that there is a blocking chain from y to z. ∎

The rest of the proof of Theorem 1 remains unchanged.

(iii) Consider a farsighted stable set F of a coalitional game such that for every $y \notin F$ there is a blocking chain from y to $x \in F$ with $\pi(x) = N$. Suppose F satisfies Property A$'$. Again, it will suffice to prove an appropriately modified version of Lemma 2 that provides a blocking chain rather than an acceptable chain.

[35]This is weaker than the corresponding assumption stated in Corollary 1, namely that $\pi(x) = N$ for every $x \in F$.

For every state $x \notin F$, fix any blocking chain with $\Psi(x) \in F$ as the terminal state such that $\pi(\Psi(x)) = N$.

LEMMA **4.** *Consider a farsighted stable set F satisfying Property A′ and condition (iii). Suppose T moves from state $x \notin F$ to state y, $\Psi(x) = a$ and $\Psi(y) = b$, with $\pi(b) = N$. 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 $j \in T$.*

Proof. Fix states x, y, a, b and a coalition T and assume that $u_T(b) \gg u_T(a)$. Of course, by farsighted internal stability, $T \neq N$, so $y \neq b$ and $y \notin F$. Fix one canonical blocking chain from y to b, say, $\{y, y^1, \ldots, y^{m-1}, y^m\}, \{S^1, \ldots, S^m\}$, where $y^m = b$. Proceeding exactly as in the proof of Lemma 2, we can find a player $j \in T$ and a new blocking chain $\{y, y^1, \ldots, y^{m-1}, y', y^m\}, \{S^1, \ldots, S^{m-1}, W - j, S^m\}$, such that at state y' player j has yet to move and $u_j(y') = 0$. Because $\pi(b) = N$, $S^m = N$: all players must be involved in the final move of the blocking chain from y to b and $u(b) \gg 0$.

Pick z as given by Property A′. We will now construct a blocking chain from y to z. (This is not surprising given Footnote 18). Consider the two cases corresponding to whether $u_j(z) > 0$:

(i) If $u_j(z) > 0$ consider the chain $\{y, y^1, \ldots, y^{m-1}, y', z\}, \{S^1, \ldots, S^{m-1}, W - j, N\}$. In the original blocking chain all players were active movers, so all players in $N - j$ are also strictly better off at z compared to the first step at which they made a move in the original blocking chain. The move to z is also strictly profitable for j. This means that we have a found a blocking chain from y to z.

(ii) If $u_j(z) = 0$ consider the chain $\{y, y^1, \ldots, y^{m-1}, y', z\}, \{S^1, \ldots, S^{m-1}, W - j, N - j\}$. The last step is feasible for $N - j$ because z is a regular state with $u_i(z) \geq u_i(b) > 0$ for all $j \neq i$ and $u_j(z) = 0$. It is also clearly a blocking chain. ■

This completes the proof of Proposition 1. ■

APPENDIX A.2. PROOF OF THEOREM 2 AND REMARK 7

Proof. The argument relies on a version of Lemma 2 that does not invoke Property B. Consider any farsighted stable set F. If $x \in F$, define $\Psi(x) = x$. For each state $x \notin F$, fix any blocking chain, $\mathbf{c}(x)$, that reaches F in a *minimal* number of steps. In particular, if there is a coalition T such that $T \in E(x, a)$ with $a \in F$ and $u_T(a) \gg u_T(x)$, then $\mathbf{c}(x)$ must reach F in one step. Let $\Psi(x) \in F$ be the terminal state for this minimal chain.

LEMMA **5.** *Let F be a farsighted stable set satisfying Property A. Suppose that in a minimal blocking chain* $\mathbf{c}$*, coalition T initiates the first move from $x \notin F$ to y'. Suppose T moves instead from x to $y \neq y'$. Let $\Psi(x) = a$ and $\Psi(y) = b$. Then there is a state $z \in F$ and an acceptable chain from y to z such that $u_j(z) \leq u_j(a)$ for some $j \in T$.*

Proof. Fix states x, y, a, b and a coalition T as in the statement of the lemma and suppose that $u_T(b) \gg u_T(a)$. If $y \in F$, because **c** is a blocking chain initiated by T that leads to F in a minimal number of steps, a must also be the result of a one-step move by coalition T, i.e., $y' = a$. Notice that all players not in T are in the same coalition in a as in b and their payoffs are also the same in a and b. The only difference between states a and b is that T may be organized in a different partition and $u_T(b) \gg u_T(a)$. Clearly, $T \in E(a, b)$, but this contradicts internal stability of F. Thus, $y \notin F$ (and so $b \neq y$). The rest of the proof is the same as that of Lemma 2. ∎

Lemma 5 can be used along with the rest of the proof of Theorem 1 to complete the proof of Theorem 2. ∎

In simple games even Property A can be dropped from Lemma 5. Fix states x, y, a, b and a coalition T as in the statement of Lemma 5 and suppose that $u_T(b) \gg u_T(a)$. In a simple game T must be a losing coalition; if it is a winning coalition, it could have gone from a to b on its own, contradicting internal stability. But a losing coalition that was supposed to move from x to y' cannot move anywhere other than y' (because of coalitional sovereignty), so in fact it has no available "deviation", and the conclusion of Lemma 5 follows. As in the proof of Proposition 1, we can strictly embed a farsighted stable set in a process satisfying maximality. It is for absolute maximality that we rely on Property A$'$. We therefore have a proof of Remark 7 as well as the following extended proposition:

PROPOSITION **2.**

(i) Every farsighted stable set of a simple game can be strictly embedded in a process satisfying maximality.

(ii) Moreover, if it satisfies Property A$'$, then it can be strictly embedded in a process satisfying absolute maximality.

Dutta and Vartiainen (2018) show that in a variant of the class of simple games the conclusion in (i) holds for strong maximality.