Rochester Center for Economic Research

Noncooperative Foundations of Stable Sets in Voting Games

Anesi, Vincent

Working Paper No. 551
May 2009

UNIVERSITY OF ROCHESTER

Noncooperative Foundations of Stable Sets in Voting Games

Vincent Anesi*
University of Nottingham

27th May 2009

Abstract

This note investigates the noncooperative foundations of von Neumann-Morgenstern (vN-M) stable sets in voting games. To do so, we study subgame perfect equilibria of a noncooperative legislative bargaining game, based on underlying simple games. The following results emerge from such an exercise: Every stable set of the underlying simple game is the limit set of undominated pure-strategy Markov perfect equilibria, and of strategically stable sets of undominated subgame perfect equilibria of the bargaining game with farsighted voters.

Keywords: Legislative bargaining, committee, strategic stability, stable set.

JEL Classification Numbers: C71, C78, D71

---

*Address: School of Economics, Room B18, The Sir Clive Granger Building, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom . Email: vincent.anesi@nottingham.ac.uk. I thank John Duggan and two anonymous referees for extremely valuable comments. I am also grateful to Debraj Ray, Daniel Seidmann, and Tasos Kalandrakis for helpful discussions and comments. The final version of this note was completed while the author was visiting the W. Allen Wallis Institute of Political Economy at the University of Rochester. He wishes to express his appreciation to the Institute for its hospitality.

1 Introduction

Simple games – also called voting games, or committees – have played a major role in the formal theory of committee voting. The key ingredients in this type of cooperative game are: (i) a set of feasible alternatives; (ii) a set of committee members (or voters); (iii) the committee members' preferences over the set of feasible alternatives; and (iv) a collection of coalitions, called *winning coalitions*, which are all-powerful in that they can rule out any alternative irrespective of the other voters' behavior. The details of the institutional setting and the distribution of power among individuals lurk in this collection of winning coalitions.

Although the concept of *stable set*, as defined by von Neumann and Morgenstern (1944), is central in the theory of cooperative games, it has so far received little attention from political scientists who study committee voting with the simple-game approach.[1] The main reason for this is the absence of informal – but credible – "stories" of individual interaction that would provide interpretations for stable sets in the context of voting games. Paraphrasing Ordeshook (1986), "the arguments for supposing that people choose outcomes in [stable] sets often remain obscure".[2] An important issue, then, concerns the extent to which stable set predictions reflect the equilibrium predictions of credible noncooperative games of negotiation and coalition formation. Put differently, given a stable set of a simple game, does there exist an institutional arrangement that implements the stable set?

The present note is concerned with developing strategic foundations for stable sets in simple games. The noncooperative framework we use to do so, legislative bargaining, is based on the above-mentioned ingredients of underlying simple games. At each stage in the negotiation, there is an initial status quo, and a committee member is given the opportunity to propose an alternative which is then voted up or down by the committee; if voted up, the alternative is implemented and becomes the new status quo; if voted down, the status quo remains unchanged and is implemented. An alternative is accepted if and only if the set of voters who accept the proposal is a wining coalition of the underlying simple game. The sequence repeats indefinitely.

We answer the question: "What reasons do we have to suppose that committees choose alternatives in stable sets?" by proving that, when voters

[1] The literature on the most familiar and widely-used solution concept, the core, is immense and will not be surveyed here. Le Breton and Weber (1992) derive necessary and sufficient conditions for the core to constitute a vN-M stable set. Anesi (2006) shows that alternatives in stable sets generalize the usual "corelike" stability notion to dynamic cooperative settings.

[2] In Ordeshook's book, stable sets are referred to as "V-sets".

are sufficiently farsighted, every stable set of the underlying simple game is the limit set of undominated subgame perfect equilibria of the noncooperative bargaining game. We further show that some of those equilibria have desirable properties, namely Markov perfection and in particular strategic stability (as defined by Kohlberg and Mertens, 1986), thus establishing a relationship between vN-M stability and strategic stability in voting games.

This short note thus contributes to the literature on noncooperative foundations of cooperative solutions in voting games. These authors have developed noncooperative games of bargaining and coalition formation to provide strategic foundations for a variety of cooperative solutions. Although this literature is now too large for us to give an exhaustive survey here, we should just mention Harsanyi (1974), who uses a different bargaining setting to provide strategic foundations for stable sets in TU games.[3] To the best of our knowledge, this note constitutes the first attempt to provide noncooperative foundations for stable sets in a voting context with nontransferable utility. More distantly related to this note is the political economy literature on legislative bargaining models with an endogenous status quo. This includes Baron (1996), Kalandrakis (2004), Battaglini and Coate (2007a,b), Battaglini and Palfrey (2007), Duggan and Kalandrakis (2007), and Diermeier and Fong (2009a,b), to name a few.

We present the underlying simple game and the corresponding legislative bargaining game in Section 2. In Section 3, we state and prove the central result of this note.

2 Notation and Definitions

2.1 The Underlying Simple Game

Let $N \equiv \{1, \ldots, n\}$ denote the set of voters (or players), indexed by i, and let $\mathcal{N} \equiv 2^N \setminus \{\emptyset\}$. Each member of $\mathcal{N}$ is referred to as a coalition. Every voter i has preferences over a finite set of alternatives X. These preferences are represented by a linear order $\succ_i$ on X. For future reference, we assume that, for every $i \in N$, $\succ_i$ is represented by a von Neumann-Morgenstern payoff function $u_i : X \to \mathbb{R}$, such that for all $(x, y) \in X^2$: $x \succ_i y$ if and only if $u_i(x) > u_i(y)$.

A *simple game* (or *voting game*) consists of a pair $(N, \mathcal{W})$, where $\mathcal{W} \subseteq \mathcal{N}$ is the collection of winning coalitions. Thus, we say that $x \in X$ *dominates* $y \in X$ if, and only if, there is a coalition $S \in \mathcal{W}$ such that $u_i(x) > u_i(y)$, for

[3] The reader is referred to Montero (2006) and all the references therein for a more comprehensive review of this literature.

every $i \in S$. Let $D(y)$ denote the set of alternatives that dominate y. We assume throughout that $\mathcal{W}$ is:

- *monotonic:* $S \in \mathcal{W}$ and $N \supseteq T \supset S \Rightarrow T \in \mathcal{W}$;
- *proper:* $S \in \mathcal{W} \Rightarrow N \setminus S \notin \mathcal{W}$.

A set of alternatives $V \subseteq X$ is a *stable set* of $(N, \mathcal{W})$ if, and only if, the two following conditions hold:

$$x \in V \Rightarrow D(x) \subseteq X \setminus V, \tag{1}$$

and

$$y \in X \setminus V \Rightarrow \exists x \in V : x \in D(y). \tag{2}$$

These two conditions are called *internal stability* and *external stability*, respectively. The existence and the uniqueness of stable sets is studied in Muto (1984).

2.2 The Bargaining Game

We now present the legislative bargaining game based on $(N, \mathcal{W})$.

Bargaining Procedure

In each of an infinite number of discrete periods, indexed $t = 1, 2, \ldots$, members of N have to collectively choose an alternative x^t from X. The sequential bargaining process is as follows: At the start of each period t, player i is chosen with probability $p_i > 0$ to make a proposal $y \in X$; once the proposal is made, all players simultaneously vote "yes" or "no". If the set of players who vote "yes" belongs to $\mathcal{W}$ then $x^t = y$; otherwise $x^t = x^{t-1}$. In each period t, x^{t-1} is thus regarded as the status quo. For simplicity, we assume that the status quo of period 1, x^0, is chosen by Nature according to probability distribution p_0 on X such that $p_0(\{x\}) > 0$ for every $x \in X$.

Once alternative x^t has been chosen, every player i receives an instantaneous payoff $(1 - \delta_i) u_i(x^t)$, where δ_i is i's discount factor. Thus, player i's payoff from a bargaining sequence $\{x^t\}_{t=1}^{\infty}$ is $(1 - \delta_i) \sum_{t=1}^{\infty} \delta_i^{t-1} u_i(x^t)$.

Strategies

A *history* at some stage of a given period t describes all that has transpired in the previous periods and stages (the sequence of proposers, their respective proposals and the associated pattern of votes). This stage may be of various kinds: a proposer is about to be selected, or a proposal about to be made, or voters about to vote, or an alternative about to be implemented. We must therefore distinguish between the corresponding types of histories: "selection histories", "proposer histories", "voter histories", and "implementation histories" respectively.

A *strategy* σ_i for a player i is a mapping that assigns a probability distribution over intended actions (what to propose, how to vote) to all conceivable proposer and voter histories. Formally, let ΔX be the family of probability distributions over X. Let H_p^t and H_v^t stand for the sets of possible proposer and voter histories of period t, respectively, and let $H \equiv \bigcup_t \left(H_p^t \cup H_v^t\right)$. For each player i, let $\pi_i^t : H_p^t \to \Delta X$ denote i's *proposal strategy* for period t (conditional on i being selected to make a proposal for t) and, for each $h \in H_p^t$ and each $x \in X$, let $\pi_i^t(h)(x)$ be the probability that proposer i makes proposal x at history h. Player i's *voting strategy* for period t is $v_i^t : H_v^t \to [0,1]$ where, for every $h \in H_v^t$, $v_i^t(h)$ is the probability that i votes "yes" at voter history h. Thus, the full collection $\sigma_i = \{(\pi_i^t, v_i^t)\}_{t=1}^{\infty}$ is a strategy for i, and $\sigma = \{\sigma_i\}_{i \in N}$ is a strategy profile.

For each $h \in H_p^t$, let $\mathbf{r}_p(h) \in X$ be the status quo prevailing at h, and for each $h \in H_v^t$, let $\mathbf{r}_v(h) \in X^2$ be the pair (x, y) where x is the ongoing status quo and y is the proposal just made at h. For future reference, player i's strategy σ_i is said to be:

– a pure strategy if and only if, for any period t: $\pi_i^t(h)(x) \in \{0,1\}$ for all $h \in H_p^t$ and all $x \in X$, and $v_i^t(h) \in \{0,1\}$ for all $h \in H_v^t$;

– a completely mixed strategy if and only if, for any period t: $\pi_i^t(h)(x) > 0$ for all $h \in H_p^t$ and all $x \in X$, and $v_i^t(h) > 0$ for all $h \in H_v^t$;

– a Markovian strategy if and only if, for all periods t and t': $\pi_i^t(h) = \pi_i^{t'}(h')$ for all $h \in H_p^t$ and $h' \in H_p^{t'}$ such that $\mathbf{r}_p(h) = \mathbf{r}_p(h')$, and $v_i^t(h) = v_i^{t'}(h')$ for all $h \in H_v^t$ and $h' \in H_v^{t'}$ such that $\mathbf{r}_v(h) = \mathbf{r}_v(h')$.

Every strategy profile σ generates a probability distribution over infinite sequences of alternatives $\{x^t\}_{t=1}^{\infty}$. Say that σ is *absorbing* if $\{x^t\}_{t=1}^{\infty}$ converges almost surely. For every period t, let $P_t^\sigma : X^2 \to [0,1]$ be the transition probability engendered by σ such that $P_t^\sigma(x,y)$ is the probability that $x^t = y$ given that $x^{t-1} = x$. The set of absorbing points of σ is then defined as

$$A(\sigma) \equiv \{x \in X : P_t^\sigma(x,x) = 1 \text{ , } \forall t = 1, \ldots, \infty\}.$$

Equilibrium

As in the previous literature, we will concentrate on undominated subgame perfect equilibria (SPEs) — namely SPEs in which no player uses a weakly dominated strategy — and be particularly interested in Markov perfect equilibria (MPEs) — those in which players use Markovian strategies. However, another equilibrium refinement will be used in the present note: Kohlberg' and Mertens' (1986) "strategic stability".[4]

A set Σ of Nash equilibria of the bargaining game is *stable* if it is minimal with respect to the following property:

(P) Σ is a closed set of equilibria satisfying: for any $\varepsilon > 0$ there exists some $\zeta_0 > 0$ such that for any completely mixed strategy profile $\{\sigma_i\}_{i \in N}$ and any numbers $\{\zeta_i\}_{i \in N}$, with $0 < \zeta_i < \zeta_0$, the perturbed game where every strategy s of player i is replaced by $(1 - \zeta_i)\, s + \zeta_i \sigma_i$ has an equilibrium that is within ε of some equilibrium in Σ.[5]

SPEs that form a stable set are consistent with both the ideas of backward and forward induction. We refer the reader to Kohlberg (1990) for an extensive discussion of the latter concept.

The name "stable set" should not cause confusion with von Neumann-Morgenstern (vN-M) stable sets as long as it is understood that a vN-M stable set refers to a set of alternatives, whereas a strategically stable set refers to a set of strategy profiles in the bargaining game. With this caveat in mind, we now turn to the implementation of vN-M stable sets in the noncooperative bargaining game.

3 Implementing a Stable Set

The main result of this note is the following proposition, which focuses on situations where voters are farsighted. Parts (i) and (ii) of the proposition show that every stable set of the underlying simple game is the absorbing set of undominated SPEs of the legislative bargaining game. Parts (iii) and (iv) tell us that something stronger is actually true: there exist a set of undominated pure-strategy MPEs and a strategically stable set of undominated SPEs that all converge to the stable set under consideration.

[4] We refer the reader to De Sinopoli (2000, 2004), and De Sinopoli and Turrini (2002) for a discussion of strategic stability in different political-economy models.

[5] Let $A_i(h)$ be the finite action set of player i at history h. That is, $A_i(h) = X$ if h is a proposer history at which i has been selected, and $A_i(h) = \{\text{Yes}, \text{No}\}$ if h is a voter history, and $A_i(h) = \emptyset$ otherwise. Thus, a strategy for player i *at a given history* h, say $\sigma_i(h)$, is an element of $\Delta A_i(h)$. The distance between two strategy profiles σ and σ' is defined as $d(\sigma, \sigma') \equiv \sup_{h \in H} d_h(\sigma(h), \sigma'(h))$, where

$$d_h(\sigma(h), \sigma'(h)) \equiv \sum_{i \in N} \sum_{a \in A_i(h)} |\sigma_i(h)(a) - \sigma_i'(h)(a)| \ .$$

Proposition 1 *Let $V \subseteq X$ be a stable set of the underlying simple game. Then there exist a set of strategy profiles, Σ^*, and $\bar{\delta} \in (0,1)$ such that the following is true whenever $\min_{i \in N} \delta_i > \bar{\delta}$:*

(i) $A(\sigma^*) = V$ for every $\sigma^* \in \Sigma^*$;

(ii) every σ^* in Σ^* is an undominated SPE;

(iii) Σ^* includes a subset of Markovian pure-strategy profiles; and

(iv) Σ^* includes a subset that is stable.

To prove Proposition 1, we first define Σ^* and $\bar{\delta}$. For every $x \notin V$, pick an arbitrary element, say $\vec{x}$, in the set $V \cap D(x)$ which, by external stability, is nonempty. For expositional convenience, we adopt the convention that $\vec{x} = x$ when $x \in V$. Note that, although there may be several functions, $x \mapsto \vec{x}$, that satisfy these conditions, only one of them is chosen and kept fixed throughout the proof of Proposition 1. Let Σ^* be the closed set of strategy profiles that satisfy the following conditions:

At the proposer history $h \in H_p^t$ of period t with status quo x (namely $x^{t-1} = x$), proposer k's strategy is the following: if $x \notin V$ and $u_k(\vec{x}) > u_k(x)$, then

$$\pi_k^t(h)(\vec{x}) = 1 \; ; \tag{3}$$

if $x \notin V$ and $u_k(\vec{x}) < u_k(x)$, then $\pi_k^t(h)$ is an element in ΔX that satisfies

$$\pi_k^t(h)(\vec{x}) = 0 \; ; \tag{4}$$

if $x \in V$, then $\pi_k^t(h)$ is some arbitrary element in ΔX:

$$\sum_{y \in X} \pi_k^t(h)(y) = 1 \; . \tag{5}$$

At the voter history $h \in H_v^t$ of period t with status quo x and proposal y, every responder i plays as follows: if $y \neq \vec{x}$, then

$$v_i^t(h) = \begin{cases} 1 & \text{if } u_i(\vec{y}) > u_i(\vec{x}) \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

and if $y = \vec{x}$, then

$$v_i^t(h) = \begin{cases} 1 & \text{if } u_i(\vec{x}) > u_i(x) \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Note that every $\sigma^* \in \Sigma^*$ induces a time-independent transition process P^{σ^*} such that $P_t^{\sigma^*} = P^{\sigma^*}$ for any period t.

We now define $\bar{\delta}$. For each player $i \in N$, and every pair $(x, y) \in X^2$ such that $u_i(\vec{y}) > u_i(\vec{x})$, let

$$\begin{aligned}
\lambda_x &\equiv \sum_{i \in N: u_i(x) > u_i(\vec{x})} p_i, \\
f_i(x) &\equiv \begin{cases} \frac{1}{1-\delta_i \lambda_x} \left[(1-\delta_i) u_i(x) + \delta_i (1-\lambda_x) u_i(\vec{x}) \right] & \text{if } x \notin V \\ u_i(x) & \text{if } x \in V. \end{cases} \\
\Lambda_i(x, y) &\equiv f_i(y) - f_i(x).
\end{aligned}$$

A brief inspection of the above definitions reveals that there exists $\bar{\delta}_i(x, y) \in (0, 1)$ such that $\Lambda_i(x, y)$ belongs to the $[u_i(\vec{y}) - u_i(\vec{x})]$-neighborhood of $u_i(\vec{y}) - u_i(\vec{x})$ (and is therefore positive) whenever $\delta_i > \bar{\delta}_i(x, y)$. We define $\bar{\delta}$ as

$$\bar{\delta} \equiv \max_{i \in N} \max_{(x,y): u_i(\vec{y}) > u_i(\vec{x})} \bar{\delta}_i(x, y).$$

We now prove that $A(\sigma^*) = V$. When the status quo, say x, is an element of V, the proposer offers a policy, say y, which is always rejected. Inspection of (6)-(7) indeed reveals that y is accepted if and only if there is a winning coalition of voters who all prefer $\vec{y}$ to $\vec{x}$. But this is impossible since $\vec{x}$ and $\vec{y}$ both belong to V which is internally stable.

Suppose now that the status quo, say x, is not a member of V. This implies that $\vec{x} \in V \cap D(x)$. Then, the coalition S of players who prefer $\vec{x}$ to x belongs to $\mathcal{W}$. If the proposer is a member of S, then she proposes $\vec{x}$ which, according to (7), is accepted. If the proposer is not a member of S, she may propose another policy. But (6) tells us that this proposal is rejected: V satisfying internal stability, the elements of V cannot dominate each other. x consequently remains the status quo until the next period involving a proposer in S. In that period, the proposer successfully proposes $\vec{x}$.

A brief observation of conditions (3)-(7) reveals that Σ^* contains Markovian pure-strategy profiles: If, at all proposer histories with status quo x, proposer k always makes the same proposal, say y_x, with a probability of 1, then the behaviors prescribed by (3)-(7) define Markovian pure strategies. This proves part (iii) of the proposition.

Let $\min_{i \in N} \delta_i > \bar{\delta}$. To establish statement (ii), we must show that every $\sigma^* \in \Sigma^*$ is an undominated SPE. We do this in two easy-to-prove steps.

Step 1: For every $i \in N$ and $x \in X$, define the value of x for voter i as

$$V_i^{\sigma^*}(x) \equiv (1-\delta_i)u_i(x) + \delta_i \sum_{z \in X} P^{\sigma^*}(x,z) V_i^{\sigma^*}(z).$$

Then, at the voter history $h \in H_v^t$ of all periods t with status quo x and proposal y, $v_i^t(h) = 1$ if and only if $V_i^{\sigma^*}(y) > V_i^{\sigma^*}(x)$.

According to (3)-(7), we can write $V_i^{\sigma^*}(\cdot)$ as follows:

$$V_i^{\sigma^*}(x) = \begin{cases} \frac{1}{1-\delta_i\lambda_x}\left[(1-\delta_i)\,u_i(x) + \delta_i\,(1-\lambda_x)\,u_i(\vec{x})\right] & \text{if } x \notin V \\ u_i(x) & \text{if } x \in V \end{cases}$$

If $y \neq \vec{x}$, then $u_i(\vec{y}) > u_i(\vec{x})$ (or equivalently $v_i^t(h) = 1$) implies that

$$V_i^{\sigma^*}(y) - V_i^{\sigma^*}(x) = \Lambda_i(x,y) > 0\ , \tag{8}$$

where the last inequality results from $\delta_i > \bar{\delta} \geq \bar{\delta}_i(x,y)$. By contrapositive, the reverse statement is also true. To see this equivalence when $y = \vec{x}$, just note that

$$V_i^{\sigma^*}(\vec{x}) - V_i^{\sigma^*}(x) = \frac{1-\delta_i}{1-\delta_i\lambda_x}\left[u_i(\vec{x}) - u_i(x)\right]. \tag{9}$$

Step 2: At the proposer history $h \in H_p^t$ of any period t starting with status quo x, proposer k cannot gain by deviating from $\pi_k^t(h)$ and conforming to σ_k^* thereafter.

Suppose first that the status quo x belongs to V. In such a case, any proposal by k is rejected by the committee. Therefore, any proposal is a best response and there is no profitable deviation from (5).

If $x \notin V$, then k must compare the value from offering $\vec{x}$, $u_k(\vec{x})$, with the value from offering any other policy y, $V_k^{\sigma^*}(x)$. Thus, it is optimal for k to propose $\vec{x}$ if $u_k(\vec{x}) > u(x)$, and any other policy otherwise. This proves that she has no profitable deviation from (3) and (4).

By the one-shot deviation principle, Steps 1 and 2 establish that every $\sigma^* \in \Sigma^*$ is an undominated SPE, thus proving part (ii) of Proposition 1. Let Σ_i^* be the set of collections $\{(\pi_i^t, v_i^t)\}_{t=1}^{\infty}$ satisfying (3)-(7) for player i. Note that, by construction, any element of Σ_i^* is a best response to any element of $\times_{j \neq i}\Sigma_j^*$, for every $i \in N$.

We now turn to part (iv). The proof is by contradiction. Fix $\varepsilon > 0$. Suppose there is a sequence ζ_0^m, with $\lim_{m\to\infty}\zeta_0^m = 0$, such that: for all m, there exists a completely mixed strategy profile $\sigma^m = (\sigma_1^m, \ldots, \sigma_n^m)$ and numbers $\zeta_1^m, \ldots, \zeta_n^m$, with $0 < \zeta_i^m < \zeta_0^m$, such that the perturbed game where every strategy s of player i is replaced by $(1-\zeta_i^m)\, s + \zeta_i^m \sigma_i^m$ has no equilibrium that is within ε of Σ^*. For future reference, let $\left\{\left(\mu_{i,t}^m, \beta_{i,t}^m\right)\right\}_{t=1}^{\infty} \equiv \sigma_i^m$, for each $i \in N$.

For every proposer history $h \in H_p^t$, let $\Pi_k^t(h)$ be the finite set of *degenerate* probability distributions, $\pi_k^t(h)$, satisfying conditions (3)-(5).[6] Similarly, for every voter history $h \in H_v^t$, let $v_i^t(h) \in \{0,1\}$ be uniquely defined by conditions (6)-(7). Then, define $\mathcal{G}^m$ as the legislative-bargaining game where the "action set" of proposer k at proposer history $h \in H_p^t$ is

$$\left\{(1-\zeta_k^m)\, \pi_k^t(h) + \zeta_k^m \mu_{k,t}^m(h) : \pi_k^t(h) \in \Pi_k^t(h)\right\} \ ,$$

and the "action set" of player i at voter history $h \in H_v^t$ is the singleton $\left\{(1-\zeta_i^m)\, v_i^t(h) + \zeta_i^m \beta_{i,t}^m(h)\right\}$.

Thus, for each m, $\mathcal{G}^m$ is an infinite-horizon game with a finite number of possible actions at each history. Fudenberg and Levine (1983) establish the existence of a SPE in such games. Therefore, $\mathcal{G}^m$ has an equilibrium, which in turn implies by construction that there exists $\hat{\sigma}^m = (\hat{\sigma}_1^m, \ldots, \hat{\sigma}_n^m) \in \Sigma^*$ such that, for every player i, $(1-\zeta_i^m)\, \hat{\sigma}_i^m + \zeta_i^m \sigma_i^m$ is a best response within $\{(1-\zeta_i^m)\, \varsigma_i + \zeta_i^m \sigma_i^m : \varsigma_i \in \Sigma_i^*\}$ to $\left\{\left(1-\zeta_j^m\right) \hat{\sigma}_j^m + \zeta_j^m \sigma_j^m\right\}_{j \neq i}$.

One can find a sufficiently large $M_1 > 0$ such that $\breve{\sigma}^m \equiv \{(1-\zeta_i^m)\, \hat{\sigma}_i^m + \zeta_i^m \sigma_i^m\}_{i \in N}$ is within ε of $\hat{\sigma}^m$ whenever $m > M_1$.[7] By assumption, the perturbed game where *every* strategy s of player j (not necessarily in Σ_j^*) is replaced by $\left(1-\zeta_j^m\right) s + \zeta_j^m \sigma_j^m$ has no equilibrium within ε of $\hat{\sigma}^m$. As a consequence, for each m, there is a strategy $\tilde{\sigma}_i^m \notin \Sigma_i^*$ such that some player i can profitably deviate from $(1-\zeta_i^m)\, \hat{\sigma}_i^m + \zeta_i^m \sigma_i^m$ to $(1-\zeta_i^m)\, \tilde{\sigma}_i^m + \zeta_i^m \sigma_i^m$. Put differently, i can gain by playing $\tilde{\sigma}_i^m \notin \Sigma^*$ instead of $\hat{\sigma}_i^m$ when every $j \neq i$ plays $\left(1-\zeta_j^m\right) \hat{\sigma}_j^m + \zeta_j^m \sigma_j^m$.

The reminder of the proof shows that this leads to a contradiction. By the one-shot deviation principle, we can restrict attention to strategies $\tilde{\sigma}_i^m$ that agree with $\hat{\sigma}_i^m$ at all histories in H (at which i is active) but one. We must distinguish between several cases:

a) $\tilde{\sigma}_i^m$ disagrees with $\hat{\sigma}_i^m$ at a proposer history of a period with status quo $x \notin V$ and $u_i(\vec{x}) > u_i(x)$. From (9), $\hat{\sigma}^m \in \Sigma^*$ implies that $V_i^{\hat{\sigma}^m}(\vec{x}) > V_i^{\hat{\sigma}^m}(x)$ for all m. As $\tilde{\sigma}_i^m$ is a strictly profitable deviation from $\hat{\sigma}_i^m$, there must be a proposal $y \neq \vec{x}$ such that i can gain by proposing y instead of $\vec{x}$. As m becomes arbitrarily large, however, $\breve{\sigma}^m$ becomes arbitrarily close to $\hat{\sigma}^m$. Since

[6] Note that $|\Pi_k^t(h)| = 1$ whenever h is a proposer history of period t with status quo x such that $x \notin V$ and $u_k(\vec{x}) > u_k(x)$.

[7] Using the notation introduced in Footnote 5, we have $d(\hat{\sigma}^m, \breve{\sigma}^m) \leq \sup_{h \in H} [\zeta_0^m d_h(\hat{\sigma}^m(h), \sigma^m(h))]$. As $d_h(\hat{\sigma}^m(h), \sigma^m(h)) < n|X|$ for all h and $\zeta_0^m \to 0$, this inequality implies that there exists M_1 such that $d(\hat{\sigma}^m, \breve{\sigma}^m) < \varepsilon$ whenever $m > M_1$.

$\tilde{\sigma}_i^m$ agrees with $\hat{\sigma}_i^m$ after i's proposal, i's continuation payoff from proposing y [resp. $\vec{x}$] becomes by continuity arbitrarily close to $V_i^{\hat{\sigma}^m}(x)$ [resp. $V_i^{\hat{\sigma}^m}(\vec{x})$] (under $\hat{\sigma}^m \in \Sigma^*$, proposal y is rejected and proposal $\vec{x}$ is accepted). As a consequence, the gain from the deviation is arbitrarily close to $V_i^{\hat{\sigma}^m}(x) - V_i^{\hat{\sigma}^m}(\vec{x}) < 0$ for arbitrarily large values of m . This is a contradiction with $\tilde{\sigma}_i^m$ being a profitable deviation for all m.

b) $\tilde{\sigma}_i^m$ disagrees with $\hat{\sigma}_i^m$ at a proposer history of a period with status quo $x \notin V$ and $u_i(\vec{x}) < u_i(x)$. From (9), $\hat{\sigma}^m \in \Sigma^*$ implies that $V_i^{\hat{\sigma}^m}(\vec{x}) < V_i^{\hat{\sigma}^m}(x)$. As $\tilde{\sigma}_i^m$ is a strictly profitable deviation from $\hat{\sigma}_i^m$ and $\tilde{\sigma}_i^m \notin \Sigma_i^*$, proposer i must be strictly better-off proposing $\vec{x}$ rather than the randomization prescribed by $\hat{\sigma}_i^m$. By the same argument as in a), however, the gain from doing so becomes arbitrarily close to $V_i^{\hat{\sigma}^m}(\vec{x}) - V_i^{\hat{\sigma}^m}(x) < 0$ as m becomes arbitrarily large. This is a contradiction with $\tilde{\sigma}_i^m$ being a profitable deviation for all m.

c) $\tilde{\sigma}_i^m$ disagrees with $\hat{\sigma}_i^m$ at a proposer history of period with a status quo belonging to V. From (5) and $\tilde{\sigma}_i^m \notin \Sigma_i^*$, this is impossible.

d) $\tilde{\sigma}_i^m$ disagrees with $\hat{\sigma}_i^m$ at a voter history. As $\tilde{\sigma}_i^m$ is supposed to be a profitable deviation from $\hat{\sigma}_i^m$ for all m, voter i must be pivotal with a positive probability, say $\gamma_i^m > 0$, at this history when every $j \neq i$ plays $\left(1 - \zeta_j^m\right)\hat{\sigma}_j^m + \zeta_j^m \sigma_j^m$ (i.e., i must belong to at least one minimal winning coalition). Let the status quo and the proposal be x and y, respectively, and suppose without loss of generality that $\hat{\sigma}_i^m$ prescribes i to vote "yes" at this history. From Step 1 above, this implies that $V_i^{\hat{\sigma}^m}(y) > V_i^{\hat{\sigma}^m}(x)$. By assumption, voter i can profitably deviate from $\hat{\sigma}_i^m$ by voting "no" instead. As $\gamma_i^m > 0$, this implies that i's expected payoff from x remaining the status quo given $\breve{\sigma}^m$, say $\breve{\chi}_{i,x}^m$, is strictly greater than her expected payoff from y being implemented given $\breve{\sigma}^m$, say $\breve{\chi}_{i,y}^m$. By the same argument as in a), however, $\breve{\chi}_{i,x}^m - \breve{\chi}_{i,y}^m$ becomes arbitrarily close to $V_i^{\hat{\sigma}^m}(x) - V_i^{\hat{\sigma}^m}(y) < 0$ as m becomes arbitrarily large; a contradiction.

This proves that Σ^* satisfies property **(P)**. Let the family of subsets of Σ^* that satisfy **(P)**, $\mathcal{S}$, be ordered by set inclusion. By Zorn's Lemma, $\mathcal{S}$ has a minimal element. This completes the proof of the proposition.

What reasons do we have to suppose that committees choose alternatives in stable sets? The answer of Proposition 1 to this question is that predictions supported by vN-M stable sets are highly consistent with those supported by undominated pure-strategy MPEs on the one hand, and by strategically stable sets of undominated SPEs of the legislative bargaining game on the other hand. Parts (i) and (ii) provide bargaining foundations for vNM stable sets with a set Σ^* of absorbing (undominated) equilibria that satisfy subgame perfection. Then parts (iii) and (iv) go further, stating that

Σ^* includes a set of pure-strategy MPEs and a strategically stable set. Two remarks are in order here. First, beyond the interpretation of vNM stable sets, part (iii) also contributes to the growing literature on the existence of pure-strategy MPEs and the characterization of equilibrium absorbing sets in infinite-horizon legislative-bargaining games with an evolving status quo.[8] The second remark concerns strategic stability. Contrary to the definitions of stable sets provided in Mertens (1989, 1991) and Hillas (1990), Kohlberg' and Mertens' (1986) definition fails to satisfy one of the requirements for strategic stability, namely backward induction. Having established subgame perfection in part (ii), it is however sufficient and more natural to look for an equilibrium refinement that captures the *other* aspects of strategic stability; Kohlberg's and Mertens' (1986) stable set does so.

A brief inspection of the proof of Proposition 1 reveals that both internal and external stabilities play a decisive role in the definition of Σ^*. We end this note with some observations about these conditions. Internal stability is in fact a necessary condition for a set of alternatives to be the absorbing set of an undominated SPE of the bargaining game.

Observation 1 *If strategy profile σ is an absorbing, undominated SPE of the bargaining game, then $A(\sigma)$ satisfies internal stability.*

Proof: Let σ be an absorbing, undominated SPE and suppose that, contrary to the above statement, $A(\sigma)$ does not satisfy internal stability. This implies that there exist two distinct policies $x, y \in A(\sigma)$ and $S \in \mathcal{W}$ such that $u_i(x) > u_i(y)$ for every $i \in S$.

After the implementation of y, proposing (successfully in an undominated SPE) x would therefore make any proposer in S strictly better-off than proposing any other policy that would not change the status quo y. This is a contradiction with σ being a SPE.

□

While combining external stability with internal stability is sufficient to obtain the properties stated in Proposition 1, external stability is not a necessary condition. Indeed, a set of alternatives that is internally but not externally stable may have those properties. A simple example shows this.

Example 1 Let $N = \{1, 2, 3, 4, 5, 6\}$ and $X = \{w, x, y, z\}$, and let $\mathcal{W}$ be the

[8] See Duggan and Kalandrakis (2007), and Diermeier and Fong (2009a,b), for instance.

set of majority coalitions. Voters' preferences over X are as follows:

$$\begin{array}{ccccccc}
x & \succ_1 & z & \succ_1 & y & \succ_1 & w \\
y & \succ_2 & z & \succ_2 & x & \succ_2 & w \\
w & \succ_3 & z & \succ_3 & y & \succ_3 & x \\
y & \succ_4 & z & \succ_4 & w & \succ_4 & x \\
w & \succ_5 & z & \succ_5 & x & \succ_5 & y \\
x & \succ_6 & z & \succ_6 & w & \succ_6 & y
\end{array}$$

Note first of all that $\{z\}$ is a stable set of $(N, \mathcal{W})$ and, by Proposition 1, is an absorbing set of strategically stable SPEs of the corresponding bargaining game for any profile of payoff functions $\{u_i(\cdot)\}_{i=1,\ldots,6}$. Consider now $\{w, x, y\}$. This set of alternative is internally stable but not externally stable (z dominates all its elements). One can easily show that, if players' payoffs satisfy:

$$\begin{aligned}
u_i(w) &> \frac{1}{3}\left[u_i(w) + u_i(x) + u_i(y)\right], \forall i \in \{4, 6\} \\
u_i(x) &> \frac{1}{3}\left[u_i(w) + u_i(x) + u_i(y)\right], \forall i \in \{2, 5\} \\
u_i(y) &> \frac{1}{3}\left[u_i(w) + u_i(x) + u_i(y)\right], \forall i \in \{1, 3\},
\end{aligned}$$

the bargaining game with recognition probabilities $p_i = 1/3$, $\forall i \in N$, has a stable set of undominated SPEs, σ, such that $A(\sigma) = \{w, x, y\}$ when $\min_i \delta_i$ is sufficiently close to 1.

□

References

Anesi, V., 2006. Committees with Farsighted Voters: A New Interpretation of Stable Sets. Social Choice and Welfare 27, 595-610.

Baron, D.P., 1996. A Dynamic Theory of Collective Goods Programs. American Political Science Review 90, 316-330.

Battaglini, M., Coate, S., 2007a. A Dynamic Theory of Public Spending, Taxation, and Debt. Forthcoming in the American Economic Review.

Battaglini, M., Coate, S., 2007b. Inefficiency in Legislative Policy-Making: A Dynamic Analysis. American Economic Review 97, 118-149.

Battaglini, M., Palfrey, T.R., 2007. The Dynamics of Distributive Politics. Northwestern University, CMS-EMS Discussion Paper 1451.

De Sinopoli, F., 2000. Sophisticated Voting and Equilibrium Refinements under Plurality Rule. Social Choice and Welfare 17, 655-672.

De Sinopoli, F., 2004. A Note on Forward Induction in a Model of Representative Democracy. Games and Economic Behavior 46, 41-54.

De Sinopoli, F., Turrini, A., 2002. A Remark on Voters' Rationality in Besley and Coate Model of Representative Democracy. Journal of Public Economic Theory 4, 163-170.

Diermeier, D., Fong, P., 2009a. Endogenous Limits on Proposal Power. Unpublished.

Diermeier, D., Fong, P., 2009b. Existence and Computation of Pure-Strategy Equilibrium in Models of Legislative Bargaining with Reconsideration. Unpublished.

Duggan, J., Kalandrakis, T., 2007. Dynamic Legislative Policy Making. W. Allen Wallis Institute Working Paper No. 45.

Fudenberg, D., Levine, D., 1983. Subgame-Perfect Equilibria of Finite- and Infinite-Horizon Games. Journal of Economic Theory 31, 251-268.

Harsanyi, J.C., 1974. An Equilibrium-Point Interpretation of Stable Sets and a Proposed Alternative Definition. Management Science 20, 1472-1495.

Hillas, J., 1990. On the Definition of the Strategic Stability of Equilibria. Econometrica 58, 1365-1390.

Kalandrakis, A., 2004. A Three-Player Dynamic Majoritarian Bargaining Game. Journal of Economic Theory 116, 294-322.

Kohlberg, E., 1990. Refinement of Nash Equilibrium: The Main Ideas. In *Game Theory and Applications,* edited by T. Ichiishi, A. Neyman and Y. Tauman. San Diego: Academic Press.

Kohlberg, E., Mertens, J.-F., 1986. On the Strategic Stability of Equilibria. Econometrica 54, 1003-1037.

Le Breton, M., Weber, S., 1992. A Note on the Core and Von Neumann-Morgenstern Solutions of Simple Games. Social Choice and Welfare 9, 57-61.

Mertens, J.-F., 1989. Stable Equilibria – A Reformulation. Part I. Definition and Basic Properties. Mathematics of Operations Research 14, 575-625.

Mertens, J.-F., 1991. Stable Equilibria – A Reformulation. Part II. Discussion of the Definition, and Further Results. Mathematics of Operations Research 16, 694-753.

Montero, M., 2006. Noncooperative Foundations of the Nucleolus in Majority Games. Games and Economic Behavior 54, 380-397.

Muto, S., 1984. Stable Sets for Simple Games with Ordinal Preferences. Journal of Operations Research Society Japan 27, 250-258.

Ordeshook, P.C., 1986. Game Theory and Political Theory. Cambridge University Press, London/New York.

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