American Economic Review 2012, 102(4): 1446–1476
http://dx.doi.org/10.1257/aer.102.4.1446

Dynamics and Stability of Constitutions, Coalitions, and Clubs[†]

By Daron Acemoglu, Georgy Egorov, and Konstantin Sonin*

In dynamic collective decision making, current decisions determine the future distribution of political power and influence future decisions. We develop a general framework to study this class of problems. Under acyclicity, we characterize dynamically stable states as functions of the initial state and obtain two general insights. First, a social arrangement is made stable by the instability of alternative arrangements that are preferred by sufficiently powerful groups. Second, efficiency-enhancing changes may be resisted because of further changes they will engender. We use this framework to analyze dynamics of political rights in a society with different types of extremist views. (*JEL* D71, D72, K10)

Consider the problem of a society choosing its constitution. Naturally, the current rewards from adopting a specific constitution will influence this decision. Yet, as long as the members of the society are forward-looking and patient, the future implications of the constitution may be even more important. For example, a constitution that encourages economic activity and benefits the majority of the population may nonetheless lead to future instability or leave room for a minority to seize political control. If so, the society—or the majority of its members—may rationally shy away from adopting such a constitution. Many problems in political economy, club theory, coalition formation, organizational economics, and industrial organization have a structure resembling this example of constitutional choice.

We develop a tractable framework for the analysis of dynamic collective decisions. Consider a society consisting of a finite number of infinitely lived individuals. It starts in a particular *state*. A state in our framework represents both economic and political arrangements. In particular, it determines stage payoffs (for example, by shaping economic allocations) and also how the society can determine its future states (e.g., which subsets of individuals can change the economic allocations and

*Acemoglu: Department of Economics, Massachusetts Institute of Technology, 50 Memorial Drive, Building E52, Room 380B, Cambridge, MA 02142-1347 (e-mail: daron@mit.edu); Egorov: Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, 2001 Sheridan Rd., Jacobs Center, Room 539, Evanston, IL 60208 (e-mail: g-egorov@kellogg.northwestern.edu); Sonin: New Economic School, 47 Nakhimovsky prosp., Office 922, Moscow, 117418, Russia (e-mail: ksonin@gmail.com). We thank three anonymous referees, John Duggan, Matthew Jackson, Robert Powell, and conference and seminar participants at the ASSA 2008 (New Orleans), the 2008 Game Theory World Congress, Berkeley, Cambridge, California Institute of Technology, Harvard, Princeton, Northwestern, Stanford, and Yale for useful comments. Acemoglu gratefully acknowledges financial support from the NSF, AFOSR, and the Canadian Institute for Advanced Research.

[†] To view additional materials, visit the article page at http://dx.doi.com/10.1257/aer.102.4.1446.

political rules; see Examples 1 and 2). Our focus is on dynamic equilibria when individuals are sufficiently forward-looking. Under natural acyclicity assumptions that rule out Condorcet-type cycles, we prove the existence and characterize the structure of (*dynamically*) *stable states*. An equilibrium is represented by a mapping ϕ which designates the dynamically stable state $\phi(s_0)$ as a function of the initial state s_0. We show that the set of dynamically stable states is largely independent of the details of agenda setting and voting protocols.

Although our main focus is the noncooperative analysis of the environment outlined above, it is both convenient and instructive to start with an axiomatic characterization of stable states. This characterization relies on the observation that sufficiently forward-looking individuals do not wish to support change toward a state (constitution) that might ultimately lead to another, less preferred state (our *stability axiom*). We also introduce two other natural axioms ensuring that individuals do not support changes that give them lower utility. We characterize the set of mappings, Φ, that are consistent with these three axioms recursively and provide conditions under which there exists a unique member of Φ (Theorem 1). We show that even when Φ is not a singleton, the sets of stable states defined by any two ϕ_1, $\phi_2 \in \Phi$, are identical.

Our main results are given in Theorem 2. Under the assumptions that (i) agents have a discount factor sufficiently close to 1, and (ii) there are (small) transaction costs from changing states, the equilibria of our dynamic game for any agenda-setting and voting protocol corresponds to some $\phi \in \Phi$. Conversely, for any $\phi \in \Phi$, there exists a protocol such that the resulting noncooperative equilibrium is represented by ϕ.

Both high discount factors and transaction costs are assumed to enable a sharp characterization of the structure of stable states, though they are also reasonable in many relevant applications.[1] The high discount factor assumption is motivated by situations in which a new state, involving a different configuration of political power, can be changed immediately by those who have power (which is itself a consequence of lack of commitment in political decisions, discussed below). We also believe that most major changes in political rules and organizational forms involve transaction costs.[2] We should add, however, that the payoff implications of these transaction costs are small in our setup precisely because the discount factor is high (and thus, in equilibrium, discounted payoffs are approximately equal to what they would have been without the transaction cost; see below).

At the center of our approach is the natural lack of *commitment* in dynamic decision-making problems—those that gain additional decision-making power as

[1] These assumptions ensure that agents compare different paths putting a sufficiently large weight on payoffs in the final state that will ultimately emerge and persist rather than on payoffs in transitory states along these paths. Since we impose relatively few restrictions on protocols and preferences (in particular, no "cardinal" comparisons between payoffs in different states), cycles in the dynamic game cannot be ruled out without sufficiently forward-looking agents and without transaction costs (see Examples 3, 4, and 5 in the online Appendix).

[2] One example illustrating the plausibility of transaction costs in the context of political change comes from the emergence of democracy, studied, among others, by Acemoglu and Robinson (2000, 2006) and Lizzeri and Persico (2004). These works assume that while commitment to policies is not feasible, political institutions, such as democracy or voting rights, cannot be immediately reversed or totally disregarded once introduced (otherwise, democracy would have no value over and above a promise to implement certain policies). The most plausible reason for this is that there are transaction costs in changing political institutions (e.g., once given, voting rights cannot be taken back without incurring some costs).

a result of a reform cannot commit to refraining from further choices that would hurt the initial set of decision makers. This lack of commitment leads to two intuitive results. First, a particular social arrangement (constitution, coalition, or club) is made stable not by the absence of a powerful set of players that prefer another alternative, but because of the absence of an alternative *stable* arrangement that is preferred by a sufficiently powerful constituency. To understand why certain social arrangements are stable, we must thus study the instabilities that changes away from these arrangements would unleash. Second, dynamically stable states can be inefficient—i.e., they may be Pareto dominated by the payoffs in another state (see Theorem 3).

Our final result, Theorem 4, provides sufficient conditions for the acyclicity assumptions (used in Theorems 1 and 2) to hold when states belong to an ordered set (e.g., a subset of $\mathbb{R}$). In particular, it shows that these results apply when (static) preferences satisfy a single-crossing property or are single-peaked (and some mild assumptions on the structure of winning coalitions are satisfied). These properties are satisfied in the majority of models of static or dynamic political economy, as illustrated by the various applications discussed in the online Appendix. Theorem 4 shows that our main results are both applicable in a wide variety of environments and typically easy to apply; also, Theorems 1 and 2 apply in a range of situations in which states do not belong to an ordered set.

Below, we provide two simple examples that illustrate main insights of our theoretical model. We start with a classic example that illustrates the tension between payoffs and political power that is present in more general form throughout our analysis. We then provide a more substantive example, to which we return in Section V.

Example 1: Consider a society that consists of two social groups, E, the elite, and M, the middle class. There are three states with different payoffs and distribution of political power: (1) absolutist monarchy a, in which E rules, with no political rights for M; (2) constitutional monarchy c, in which M has greater security and is willing to invest; (3) democracy d, where M becomes more influential and privileges of E disappear. Stage payoffs satisfy

$$w_E(d) < w_E(a) < w_E(c), \text{ and } w_M(a) < w_M(c) < w_M(d).$$

This implies that E has a higher payoff under constitutional monarchy than under absolutist monarchy (e.g., because greater investments by M increase tax revenues). On the other hand, M prefers democracy to constitutional monarchy and is least well-off under absolutist monarchy. Both parties discount the stage payoffs at rate $\beta \in (0,1)$. States a, c, and d not only determine payoffs, but also specify decision rules. In absolutist monarchy, E decides which regime will prevail tomorrow; in both c and d, M decides next period's regime.

Using our notation, d is a dynamically stable state, and $\phi(d) = d$. In contrast, c is not a dynamically stable state, since starting from c, there will be a transition to d and thus, $\phi(c) = d$. Therefore, if, starting in state a, E chooses a transition to c, this will lead to d in the following period, and thus give E a discounted payoff of

$$U_E(reform) = w_E(c) + \beta \frac{w_E(d)}{1 - \beta}.$$

If E decides to stay in a forever, its payoff is $U_E(no\ reform) = w_E(a)/(1 - \beta)$. If β is sufficiently small, then $U_E(no\ reform) < U_E(reform)$, and reform takes place. When players are sufficiently forward-looking, however (β is large), then $U_E(no\ reform) > U_E(reform)$. In this case, $\phi(a) = a$. This example illustrates both of our main results. First, state a is made stable by the instability of another state, c, which is preferred by those who are powerful in a. Second, both E and M would be strictly better off in c than in a, so the stable state starting from a is Pareto inefficient. It also illustrates that the set of stable states is larger when players are forward-looking (when β is small, only d is stable; when β is large, both a and d are stable).

Example 2: Consider the choice of how inclusive society should be towards different political and social views. A central issue facing most countries with significant Muslim populations is what types of political, social, and economic rights to give to religious and secular groups. At one end, countries such as Saudi Arabia and Iran deprive secular groups of all kinds of social and legal rights. At the other end, Turkey, Syria, Algeria, and several European countries with Muslim minorities have at times restricted participation of religious individuals in political and social life. Both types of bans appear to be motivated, at least in part, by dynamic considerations. Saudi Arabia and Iran are concerned that giving rights to nonreligious groups would weaken their regimes, while in Turkey bans on Islamist practices and parties have been motivated by the so-called "slippery slope" argument that giving rights to religious groups would ultimately reduce the rights of secular groups.[3] Some commentators interpret the developments in Turkey following greater inclusiveness toward religious groups and parties as supporting the predictions of this slippery slope argument.

To capture these issues in the simplest possible way, consider a society consisting of N individuals ranked in ascending order of religiosity. A state s consists of the set of individuals Z who currently have the right to political participation and a policy ρ that determines tolerance to secularism and religiosity. Individuals receive utility from their income and from policy ρ. Suppose that the larger is the set of individuals with the right to political participation, the greater are net incomes (e.g., because the society functions more cooperatively or individuals with rights feel more secure and undertake greater investments or are less likely to rebel). We fix a political rule, e.g., majority or supermajority rule, that determines who can choose both ρ and the set of individuals who will have the right to political participation in the next period.

This is a highly complex and, in our view, interesting social situation. It captures the "slippery slope" argument that giving rights to previously excluded religious individuals has short-run economic benefits but could later deprive secular individuals of their political rights. Moreover, both the high discount factor and transaction costs appear plausible in this context.[4] In Section V, we apply our general results to the study of this environment.

[3] On "slippery slope" arguments, see Schauer (1985), and on the conflict between religious and secular groups, see Rabasa and Larrabee (2008) and Roy (2007).

[4] For example, in Turkey the first religious local administration in Istanbul quickly moved to restrict the ability of certain restaurants to serve alcohol (though ultimately the most extreme measures were not successful), which is consistent with frequent choices of actions and thus high discount factors. Furthermore, even minor constitutional changes led to significant conflict and gridlocks, with potential economic and social costs, which is consistent with significant transaction costs.

This example also enables us to investigate the question: can we change the constitution so as to give the right to political participation while at the same time ban certain policies and certain future constitutional changes? This issue can be analyzed within our framework by introducing constitutions that require unanimity for certain types of changes (see also Barberà and Jackson 2004). Such constitutions guarantee Pareto efficiency. Our analysis, however, highlights the reasons why constitutions that stipulate such unanimity rules may not be credible; e.g., when a certain supermajority has sufficient de facto political power to challenge the unanimity clause.

Roberts (1999) and Barberà, Maschler, and Shalev (2001) can be viewed as major precursors to our paper. Roberts (1999) studies dynamic voting in clubs in a society with N individuals, where voting is by majority rule, individuals are ordered according to "single-crossing" preferences, and only clubs of the form $\{1, 2, \ldots, k\}$ for different values of k are allowed. Barberà, Maschler, and Shalev (2001) study a dynamic game of club formation in which any member of the club can admit a new agent unilaterally.[5] Lagunoff (2009), who constructs a general model of political reform and relates reform to the time inconsistency of induced social rules, is another precursor. The analyses of franchise extension in Acemoglu and Robinson (2000, 2006) and Lizzeri and Persico (2004), and the model of constitutional stability in Barberà and Jackson (2004), are also related and can be cast as applications of our general framework.

Two other closely related papers are Chwe (1994) and Gomes and Jehiel (2005). Chwe provides a model where payoffs are determined by states and transitions from one state to another are governed by exogenous rules to analyze the relationship between two distinct notions from cooperative game theory, consistent and stable sets. In Chwe's setup, however, neither a noncooperative analysis nor characterization results is possible.[6] Gomes and Jehiel study a related environment with side payments. They show that a player may sacrifice his instantaneous payoff to improve his bargaining position for the future, and that the equilibrium may be inefficient when the discount factor is small. In contrast, in our game Pareto dominated outcomes are not only possible in general, but may emerge as unique equilibria and are more likely when discount factors are close to 1. We also provide a full set of characterization (and uniqueness) results, which are not present in Gomes and Jehiel (and in fact, with side payments, we suspect that such results are not possible). Finally, in our paper a dynamically stable state depends on the initial state, while in Gomes and Jehiel, as the discount factor tends to 1, there is "ergodicity" (i.e., the ultimate distribution of states does not depend on the initial state).

Finally, our work is also related to the literatures on noncooperative coalition formation and club theory.[7] An important difference between our approach and the previous literature on coalition formation is that, motivated by political settings, we assume that the majority (or supermajority) of the members of the society can impose their will on

[5] Barberà, Sonnenschein, and Zhou (1991) study a model of voting by quotas, so that a club admits a new member if sufficiently many current members (more than the quota) vote in favor. This implies that there may be many outcomes of voting at a given voting stage, while our assumptions impose that, at each voting stage, there is always a unique status quo and a unique alternative.

[6] The link between Chwe's consistent sets and our dynamically stable states is discussed in the online Appendix.

[7] On noncooperative coalition formation, see, e.g., Mariotti (1997); Ray and Vohra (1997, 1999); Seidmann and Winter (1998); and Konishi and Ray (2003). On club theory, see Ellickson et al. (1999) and Scotchmer (2002).

those players who are not a part of the majority. This contrasts with the positive externalities and free-rider problems studied by the previous literature. In addition, most of these works assume the possibility of binding commitments (Ray and Vohra 1997, 1999), while we suppose that players have no commitment power.[8]

The rest of the paper is organized as follows. Section I introduces the general environment. Section II presents our axiomatic analysis. In Section III, we prove the existence of a (pure-strategy) Markov perfect equilibrium of the dynamic game for any agenda setting and voting protocol and establish the equivalence between these equilibria and the axiomatic characterization in Section II. Section IV applies our results when states belong to an ordered set, while Section V uses our results to study the dynamics of political rights discussed in Example 2. Section VI concludes. Appendix A contains main proofs; additional proofs, applications, and examples are presented in the online Appendix.

I. Environment

There is a finite set of players $\mathcal{I}$. Time is discrete and infinite, indexed by $t\,(t \geq 1)$. There is a finite set of *states* which we denote by $\mathcal{S}$. Throughout the paper, $|X|$ denotes the number of elements of set X, so $|\mathcal{I}|$ and $|\mathcal{S}|$ denote the number of individuals and states, respectively. States represent both different institutions affecting players' payoffs, and the distribution of political power and the procedures for decision making (e.g., sizes and identities of ruling coalitions, the degree of supermajority, or the weights or powers of different agents). Although our game is one of nontransferable utility, a limited amount of transfers can be incorporated by allowing multiple (but still a finite number of) states that have the same procedure for decision making, but different payoffs across players.

The initial state is denoted by $s_0 \in \mathcal{S}$. This state may be a part of the description of the game or chosen by Nature from $\mathcal{S}$ at random. For any $t \geq 1$, the state $s_t \in \mathcal{S}$ is determined endogenously. A nonempty set $X \subset \mathcal{I}$ is called a *coalition*, and we denote the set of coalitions by $\mathcal{C}$. Each state $s \in \mathcal{S}$ is characterized by a pair $\left(\{w_i(s)\}_{i\in\mathcal{I}}, \mathcal{W}_s\right)$. Here, for each state $s \in \mathcal{S}$, $w_i(s)$ is a (strictly) positive stage payoff assigned to individual $i \in \mathcal{I}$. Political institutions in state s are described by the set of *winning coalitions* in state s, $\mathcal{W}_s$, a (possibly empty) subset of $\mathcal{C}$. This allows us to summarize different political procedures, such as weighted majority or supermajority rules, in an economical way. For example, if in state s a majority is required for decision making, $\mathcal{W}_s$ includes all subsets of $\mathcal{I}$ that form a majority; if in state s individual i is a dictator, $\mathcal{W}_s$ contains all coalitions that include i.[9] Since $\mathcal{W}_s$ is a function of the state, the procedure for decision making can vary across states.[10]

[8] Other related works include Burkart and Wallner (2000), who develop an incomplete contracts theory of club enlargement; Jehiel and Scotchmer (2001), who show that the requirement of a majority consent for admission to a jurisdiction may be no more restrictive than an unrestricted right to migrate; Alesina, Angeloni, and Etro (2005), who study the problem of EU enlargement; and Bordignon and Brusco (2003), who study the role of "enhanced cooperation agreements" in EU enlargement.

[9] Political rules summarized by the $\mathcal{W}_s$s do not specify certain institutional details, such as who makes proposals, how voting takes place, and so on. These are specified by the agenda-setting and voting protocols of our dynamic game. We will show that these only have a limited effect on equilibrium outcomes, justifying our focus on $\mathcal{W}_s$ as a representation of "political rules."

[10] Our environment allows for the case where some states, say s and s', provide the same payoffs for all players but have different sets of winning coalitions.

Throughout the paper, we maintain the following assumption.

ASSUMPTION 1 (Winning Coalitions): *For any state* $s \in \mathcal{S}$, $\mathcal{W}_s \subset \mathcal{C}$ *satisfies*

(i) *If* $X, Y \in \mathcal{C}$, $X \subset Y$, *and* $X \in \mathcal{W}_s$ *then* $Y \in \mathcal{W}_s$.

(ii) *If* $X, Y \in \mathcal{W}_s$, *then* $X \cap Y \neq \varnothing$.

Part (i) simply states that if some coalition X is winning in state s, then increasing the size of the coalition will not reverse this. Part (ii) rules out the possibility that two disjoint coalitions are winning in the same state. If $\mathcal{W}_s = \varnothing$, state s is *exogenously stable*. None of our existence or characterization results depend on whether there is an exogenously stable state.

We introduce the following binary relations on $\mathcal{S}$. For $x, y \in \mathcal{S}$, we write

$$x \sim y \Leftrightarrow \forall i \in \mathcal{I} : w_i(x) = w_i(y). \tag{1}$$

In this case we call states x and y *payoff-equivalent*, or simply *equivalent*. More important for our purposes is the binary relation $\succeq_z$. For any $z \in \mathcal{S}$, $\succeq_z$ is defined by

$$y \succeq_z x \Leftrightarrow \{i \in \mathcal{I} : w_i(y) \geq w_i(x)\} \in \mathcal{W}_z. \tag{2}$$

Intuitively, $y \succeq_z x$ means that there exists a coalition of players that is winning (in z) with each of its members weakly preferring y to x. Note three important features about $\succeq_z$. First, it contains information about stage payoffs only. In particular, $w_i(y) \geq w_i(x)$ does *not* mean that individual i prefers a switch to state y rather than x. Whether or not he does so depends on the continuation payoffs following such a switch. Second, the relation $\succeq_z$ does not presume any type of coordination or collective decision making among the members of the coalition in question. It simply records the existence of such a coalition. Third, the relation $\succeq_z$ is conditioned on z since whether the coalition of players weakly preferring y to x is winning depends on the set of winning coalitions, which is state dependent. With a slight abuse of terminology, if equation (2) holds, we say that y is *weakly preferred* to x in z. In light of the preceding comments, this neither means that all individuals prefer y to x, nor that there will necessarily be a transition from state x to y—it simply designates that there exists a winning coalition of players, each obtaining a greater stage payoff in y than in x. Relation $\succ_z$ is defined similarly by

$$y \succ_z x \Leftrightarrow \{i \in \mathcal{I} : w_i(y) > w_i(x)\} \in \mathcal{W}_z. \tag{3}$$

If (3) holds, we say that y is *strictly preferred* to x in z.[11]

[11] Relation $\sim$ defines equivalence classes; if $x \sim y$ and $y \sim z$, then $x \sim z$. In contrast, the binary relations $\succeq_z$ and $\succ_z$ need not even be transitive. Nevertheless, for any $x, z \in \mathcal{S}$, we have $x \not\succ_z x$, and whenever $\mathcal{W}_z$ is nonempty, we also have $x \succeq_z x$. From Assumption 1 we have that for any $x, y, z \in \mathcal{S}$, $y \succ_z x$ implies $x \not\succeq_z y$, and similarly $y \succeq_z x$ implies $x \not\succ_z y$.

The next assumption puts some joint restrictions on payoff functions and winning coalitions.

ASSUMPTION 2 (Preferences): *Payoffs* $\{w_i(s)\}_{i\in\mathcal{I},s\in\mathcal{S}}$ *satisfy the following properties:*

(i) *For any sequence of states* $s_1, s_2, \ldots, s_k$ *in* $\mathcal{S}$,

$$s_{j+1} \succ_{s_j} s_j \text{ for all } 1 \le j \le k-1 \Rightarrow s_1 \not\succ_{s_k} s_k.$$

(ii) *For any sequence of states* $s, s_1, \ldots, s_k$ *in* S *with* $s_j \succ_s s$ *for* $1 \le j \le k$ *and* $s_j \nsim s_l$ for $1 \le j < l \le k$,

$$s_{j+1} \succeq_s s_j \text{ for all } 1 \le j \le k-1 \Rightarrow s_1 \not\succeq_s s_k.$$

Moreover, if for $x, y, s \in S$ *we have* $x \succ_s s$ *and* $y \succ_s x$, *then* $y \succ_s s$.

Assumption 2 plays a major role in our analysis and ensures "acyclicity" (but is weaker than "transitivity"). Part (i) rules out cycles of the form $y \succ_x x$, $z \succ_y y$, $x \succ_z z$—that is, a cycle such that in each state, a winning coalition of players strictly prefers the next state. Part (ii) rules out cycles of the form $y \succeq_s x$, $x \succeq_s z$, $z \succeq_s y$ (unless the states x, y, and z are payoff-equivalent). As such, it also rules out any cycles of the form $y \succ_s x$, $z \succ_s y$, $x \succ_s z$.[12] It also imposes an additional requirement that may be interpreted as "partial transitivity."[13]

Although Assumptions 1 and 2 rule out several interesting environments, they are natural given our interest in obtaining general characterization results. More importantly, they are satisfied in most dynamic political economy models (see Theorem 4 and applications discussed in the online Appendix). In addition to Assumptions 1 and 2, we obtain additional uniqueness results by imposing the following (stronger) requirement.

ASSUMPTION 3 (Comparability): *For* $x, y, s \in \mathcal{S}$ *such that* $x \succ_s s$, $y \succ_s s$, *and* $x \nsim y$, *either* $y \succ_s x$ *or* $x \succ_s y$.

Assumption 3 means that if two states x and y are strictly preferred to s (in s), and they are not equivalent, then x and y are $\succ_s$-comparable. This assumption is not necessary for our main results but is sufficient to guarantee uniqueness of equilibrium.

II. Axiomatic Characterization

Before specifying the details of agenda-setting and voting protocols, we provide an abstract characterization of stable states. This axiomatic analysis has two

[12] Neither part of Assumption 2 is implied by the other. Examples 6 and 7 in the online Appendix illustrate the types of cycles that can arise when either 2(i) or 2(ii) fails.

[13] Transitivity would require that for any s, x, y and z, $y \succ_s x$, $x \succ_s z$ implies $y \succ_s z$. Instead, our condition imposes this only when $z = s$.

purposes. First, it illustrates that the key economic forces that arise in the context of dynamic collective decision making are largely independent of the details of agenda-setting and voting protocols. Second, the results in this section are a preparation for the characterization of the equilibrium of the dynamic game introduced in the next section. In particular, our main result, Theorem 2, will make use of this axiomatic characterization.

The key economic insight enabling an axiomatic characterization is that *with sufficiently forward-looking behavior, an individual should not wish to transit to a state that will ultimately lead to another state that gives her lower utility.* This basic insight enables a tight characterization of (*axiomatically*) *stable states*.

More formally, our axiomatic characterization determines a set of mappings Φ such that for any $\phi \in \Phi$, $\phi : \mathcal{S} \to \mathcal{S}$ assigns an axiomatically stable state $s^{\infty} \in \mathcal{S}$ to each initial state $s_0 \in \mathcal{S}$. We impose the following three natural axioms on ϕ.

AXIOM 1 (Desirability): *If $x, y \in \mathcal{S}$ are such that $y = \phi(x)$, then either $y = x$ or $y \succ_x x$.*

AXIOM 2 (Stability): *If $x, y \in \mathcal{S}$ are such that $y = \phi(x)$, then $y = \phi(y)$.*

AXIOM 3 (Rationality): *If $x, y, z \in \mathcal{S}$ are such that $z \succ_x x$, $z = \phi(z)$, and $z \succ_x y$, then $y \neq \phi(x)$.*

All three axioms are natural in light of what we have discussed above. Axiom 1 requires that the society should not permanently move from state x to another state y unless there is a winning coalition that supports this transition. Axiom 2 encapsulates the stability notion discussed above; if some state is not dynamically stable, it cannot be the ultimate stable state for any initial state. Axiom 3 imposes the reasonable requirement that if there exists a stable state z preferred to both x and y by winning coalitions in state x, then ϕ should not pick y in x.[14] Note that while all three axioms refer to properties of ϕ, they are closely related to underlying individual preferences that ϕ aggregates.

We next define the set Φ formally and state the relationship between axiomatically stable states and Φ.

DEFINITION 1 (Axiomatically Stable States): *Let* $\Phi \equiv \{\phi : \mathcal{S} \to \mathcal{S}$: *$\phi$ satisfies* Axioms 1–3$\}$. *A state $s \in \mathcal{S}$ is* (**axiomatically**) **stable** *if $\phi(s) = s$ for some $\phi \in \Phi$. The set of stable states (fixed points) for mapping $\phi \in \Phi$ is $\mathcal{D}_{\phi} = \{s \in \mathcal{S}: \phi(s) = s\}$ and the set of all stable states is $\mathcal{D} = \{s \in \mathcal{S}: \phi(s) = s$ for some $\phi \in \Phi\}$.*

The next theorem establishes the existence of stable states and paves the way for Theorem 2, which establishes the equivalence between equilibria of the dynamic

[14] Assumption 2(ii) guarantees that if $y \succ_x x$ and $z \succ_x y$, then $z \succ_x x$. Thus, if Axiom 1 is satisfied, then the requirement $z \succ_x x$ may be dropped in Axiom 3.

game (defined in Section III below) and stable sets of mappings $\phi \in \Phi$. A proof of Theorem 1 is provided in Appendix A.[15]

THEOREM 1 (Axiomatic Characterization of Stable States): *Suppose Assumptions* 1 *and* 2 *hold. Then*

(i) *The set* Φ *is nonempty. That is, there exists a mapping* ϕ *satisfying Axioms* 1–3.

(ii) *Any* $\phi \in \Phi$ *can be recursively constructed as follows. Order the states as* $\{\mu_1, \ldots, \mu_{|\mathcal{S}|}\}$ *such that for any* $1 \leq j < l \leq |\mathcal{S}|$, $\mu_l \not\succ_{\mu_j} \mu_j$. *Let* $\phi(\mu_1) = \mu_1$. *For each* $k = 2, \ldots, |\mathcal{S}|$, *define*

$$\mathcal{M}_k = \{s \in \{\mu_1, \ldots, \mu_{k-1}\} : s \succ_{\mu_k} \mu_k \text{ and } \phi(s) = s\}. \tag{4}$$

Then

$$\phi(\mu_k) = \begin{cases} \mu_k & \text{if } \mathcal{M}_k = \varnothing \\ s \in \mathcal{M}_k \colon \nexists z \in \mathcal{M}_k \text{ with } z \succ_{\mu_k} s & \text{if } \mathcal{M}_k \neq \varnothing \end{cases}. \tag{5}$$

(*If there exist more than one* $s \in \mathcal{M}_k$ *such that* $\nexists z \in \mathcal{M}_k$ *with* $z \succ_{\mu_k} s$, *pick any of these; this corresponds to multiple* ϕ *functions*).

(*iii*) *The stable sets of any two mappings* $\phi_1, \phi_2 \in \Phi$ *coincide; i.e.,* $\mathcal{D}_{\phi_1} = \mathcal{D}_{\phi_2} = \mathcal{D}$.

(*iv*) *If, in addition, Assumption* 3 *holds, then for any two mappings* ϕ_1 *and* ϕ_2 *in* Φ, $\phi_1(s) \sim \phi_2(s)$ *for all* $s \in \mathcal{S}$.

Theorem 1 provides a simple recursive characterization of the set of mappings Φ that satisfy Axioms 1–3. Intuitively, Assumption 2(i) ensures that there exists some state $\mu_1 \in \mathcal{S}$ such that there does not exist another $s \in \mathcal{S}$ with $s \succ_{\mu_1} \mu_1$. Taking μ_1 as base, we order the states as $\{\mu_1, \ldots, \mu_{|\mathcal{S}|}\}$ according to relation $\not\succ_{\mu_j}$ as indicated in part 2 of the theorem. Then, we recursively construct the set of states $\mathcal{M}_k \subset \mathcal{S}$, $k = 2, \ldots, |\mathcal{S}|$, that includes stable states that are preferred to state μ_k (that is, states s such that $\phi(s) = s$ and $s \succ_{\mu_k} \mu_k$). When the set $\mathcal{M}_k$ is empty, there exists no stable state that is preferred to μ_k (in μ_k) by members of a winning coalition. In this case, we have $\phi(\mu_k) = \mu_k$. When $\mathcal{M}_k$ is nonempty, there exists such a stable state and thus $\phi(\mu_k) = s$ for some such s. In addition to its recursive (and thus easy-to-construct) nature, this characterization is useful as it highlights the fundamental property of stable states emphasized in the Introduction: a state μ_k is made stable precisely by

[15] This theorem may be proved under weaker assumptions. Part (ii) of Assumption 2 may be substituted by the following condition: *For any sequence of states* $s, s_1, \ldots, s_k$ *in* S *with* $s_j \succ_s s$ *for* $1 \leq j \leq k$, $s_{j+1} \succ_s s_j$ *for all* $1 \leq j \leq k-1$ *would imply* $s_1 \not\succ_s s_k$.

the absence of winning coalitions in μ_k favoring a transition to another *stable* state (i.e., by the fact that $\mathcal{M}_k = \varnothing$). This insight plays an important role in applications.

Part (iii) of Theorem 1 shows that the set of stable states $\mathcal{D}$ does not depend on the specific ϕ chosen from Φ. For two different maps ϕ_1 and ϕ_2 in Φ, it is possible that $\phi_1(s_0) \neq \phi_2(s_0)$ for some initial state s_0, but the ranges of these mappings are the same. These ranges, and thus the set of stable states $\mathcal{D}$, are uniquely determined by preferences and the structure of winning coalitions.[16] Finally, part 4 shows that when Assumption 3 holds, any stable states resulting from an initial state must be payoff-equivalent. In other words, if $s_1 = \phi_1(s_0)$ and $s_2 = \phi(s_0)$, then s_1 and s_2 might differ in terms of the structure of winning coalitions, but they must give the same payoffs to each individual.

We have motivated the analysis leading up to Theorem 1 with the argument that, when agents are sufficiently forward-looking, only axiomatically stable states should be observed (at least in the "long run"; i.e., for $t \geq T$ for some finite T). The analysis of the dynamic game introduced in the next section substantiates this interpretation.

III. Noncooperative Foundations of Dynamically Stable States

We now describe the extensive-form game capturing dynamic interactions in the environment of Section I and characterize the Markov perfect equilibria (MPE) of this game. The main result is the equivalence between the MPE of this game and the set Φ in Theorem 1.

We first specify preferences and introduce transaction costs of changing states. At each date t, individual i maximizes discounted utility

$$U_i(t) = (1 - \beta)\sum_{\tau=t}^{\infty} \beta^{\tau-t} u_i(\tau), \tag{6}$$

where $\beta \in (0,1)$ is a common discount factor. We also impose

ASSUMPTION 4 (Payoffs): *The stage payoffs in* (6) *are given by*

$$u_i(t) = \begin{cases} w_i(s_t) & \text{if } s_t = s_{t-1} \\ \tilde{w}_i & \text{if } s_t \neq s_{t-1} \end{cases}. \tag{7}$$

For each $i \in \mathcal{I}$ and any state $x \in \mathcal{S}$, we have

$$\tilde{w}_i < w_i(x).$$

Assumption 4 introduces a "transaction cost" of state transitions: in any period in which there is a transition, each player obtains a lower payoff than she would

[16] In the online Appendix, we relate the set $\mathcal{D}$ to two concepts from cooperative game theory, von Neumann-Morgenstern's stable set and Chwe's largest consistent set. Under Assumptions 1 and 2, both sets coincide with $\mathcal{D}$.

have done without the transition. Given our normalization $w_i(s) > 0$, Assumption 4 is satisfied; e.g., if $\tilde{w}_i = 0$ for all i. Since we focus on the case of β close to 1, this transaction cost has little effect on discounted payoffs.[17] In particular, once (and if) a dynamically stable state s is reached, individuals receive $w_i(s)$ at each date thereafter. Substantively, this transaction cost is introduced to guarantee the existence of a pure-strategy MPE.[18]

We next specify: (i) a protocol for a sequence of agenda-setters and proposals in each state; and (ii) a protocol for voting over proposals. Voting is sequential and is described below; the exact sequence in which votes are cast will not matter.[19] We represent the protocol for agenda setting using a sequence of mappings, $\{\pi_s\}_{s\in\mathcal{S}}$, and refer to it simply as a *protocol*. Let K_s be a natural number for each $s \in \mathcal{S}$. Then, π_s is defined as a mapping

$$\pi_s : \{1, \ldots, K_s\} \to \mathcal{I} \cup \mathcal{S}$$

for each state $s \in \mathcal{S}$. Thus, each π_s specifies a finite sequence of elements from $\mathcal{I} \cup \mathcal{S}$, and determines the sequence of agenda-setters and proposals (here K_s is the length of this sequence for state s). If $\pi_s(k) \in \mathcal{I}$, then it denotes an agenda-setter who will make a proposal from the set of states $\mathcal{S}$. Alternatively, if $\pi_s(k) \in \mathcal{S}$, then it directly corresponds to an exogenously specified proposal over which individuals vote. Therefore, the extensive-form game is general enough to include both proposals for a change to a new state initiated by agenda-setters and exogenous proposals. We make the following assumption on $\{\pi_s\}_{s\in\mathcal{S}}$:

ASSUMPTION 5 (Protocols): *For each $s \in \mathcal{S}$, one (or both) of the following two conditions holds:*

(i) *For any state $z \in \mathcal{S}\backslash\{s\}$, there exists $k : 1 \leq k \leq K_s$ such that $\pi_s(k) = z$.*

(ii) *For any player $i \in \mathcal{I}$ there exists $k : 1 \leq k \leq K_s$ such that $\pi_s(k) = i$.*

This assumption implies that either sequence π_s contains all possible states other than the "status quo" s as proposals or it allows all possible agenda-setters to eventually make a proposal before the voting round ends. We assume that protocol π_s is fixed for each state s; different states might have the same payoffs and winning coalitions under different protocols.

[17] More precisely, define $\overline{\varepsilon} = \max_{i\in\mathcal{I},x\in\mathcal{S}} |w_i(x) - \tilde{w}_i|$, which is a natural measure of the size of transaction costs. Then for any $\overline{\varepsilon}$, there exists $\beta_0 < 1$ such that Theorem 2 holds for $\beta > \beta_0$. This fact, which is proved in the online Appendix, implies that payoffs from the game considered here are arbitrarily close to an environment without transaction costs.

[18] Examples 4 and 5 in the online Appendix demonstrate that if the transaction cost is removed from equation (7), a (pure-strategy) equilibrium may fail to exist or may include cycles. While these possibilities are potentially interesting, they appear to be nonrobust. Alternative game forms (e.g., those that assume a small cost of voting) lead to results similar to what we derive with the current specification.

[19] The assumption of sequential voting allows us to focus on MPE without further refinements that are typically required to rule out counterintuitive voting equilibria. Acemoglu, Egorov, and Sonin (2009) suggest an equilibrium refinement, Markov trembling hand perfect equilibrium, which implies identical equilibrium behavior for games with simultaneous voting and corresponding games with sequential voting.

In the beginning, at $t = 0$, state $s_0 \in \mathcal{S}$ is determined (either as part of the description of the environment or randomly). Subsequently (for $t \geq 1$), the timing of events is as follows:

- Period t begins with state s_{t-1} inherited from the previous period.
- For $k = 1, \ldots, K_{s_{t-1}}$, the kth proposal $P_{k,t}$ is determined as follows. If $\pi_{s_{t-1}}(k) \in \mathcal{S}$, then $P_{k,t} = \pi_{s_{t-1}}(k)$. If $\pi_{s_{t-1}}(k) \in \mathcal{I}$, then player $\pi_{s_{t-1}}(k)$ chooses $P_{k,t} \in \mathcal{S}$.
- If $P_{k,t} \neq s_{t-1}$, then there is sequential voting between $P_{k,t}$ and s_{t-1} (we will show that the sequence of voters has no effect on the equilibrium outcome). Each player votes *yes* (for $P_{k,t}$) or *no* (for s_{t-1}). Let $Y_{k,t}$ denote the set of players who voted *yes*. If $Y_{k,t} \in \mathcal{W}_{s_{t-1}}$, then alternative $P_{k,t}$ is accepted; otherwise (if $Y_{k,t} \notin \mathcal{W}_{s_{t-1}}$), it is rejected. If $P_{k,t} = s_{t-1}$, there is no voting and we adopt the convention that in this case $P_{k,t}$ is rejected.
- If $P_{k,t}$ is accepted, then a transition to state $s_t = P_{k,t}$ takes place, and the period ends. If $P_{k,t}$ is rejected or if there is no voting because $P_{k,t} = s_{t-1}$ and $k < K_{s_{t-1}}$, then the game moves to step 2 with k increased by 1; if $k = K_{s_{t-1}}$, the next state is $s_t = s_{t-1}$, and the period ends.
- At the end of the period, each player receives stage payoff $u_i(t)$.

An MPE is defined in the standard fashion as a subgame perfect equilibrium (SPE) where strategies are functions of "payoff-relevant states" only. Here payoff-relevant states are different from the states $s \in \mathcal{S}$ described above, since the proposal under consideration, as well as votes already cast, are also payoff-relevant for the continuation game (see the online Appendix for a formal definition). Any Markovian strategy profile σ in the dynamic game defines a transition mapping on $\mathcal{S}$, $s \mapsto s^\sigma$, where $s_t = s_{t-1}^\sigma$ is the next period's state given state s_{t-1}. In what follows, we use the terms MPE and equilibrium interchangeably. Next, we define dynamically stable states.

DEFINITION 2 (Dynamically Stable States): *State $s^\infty \in \mathcal{S}$ is a* **dynamically stable state** *if there exist an initial state $s_0 \in \mathcal{S}$, a set of protocols $\{\pi_s\}_{s \in \mathcal{S}}$, an MPE strategy profile σ, and $T < \infty$ such that along the equilibrium path we have $s_t = s^\infty$ for all $t \geq T$.*

Put differently, s^∞ is a dynamically stable state if it is reached in some finite time T and is repeated thereafter—$s_t = s^\infty$ for all $t \geq T$. Our objective is (i) to determine whether dynamically stable states exist in the dynamic game described above and to characterize them as a function of the initial state $s_0 \in \mathcal{S}$, and (ii) to establish the equivalence between dynamically and axiomatically stable states characterized in the previous section.

We consider situations in which β is greater than some threshold $\beta_0 \in (0,1)$ derived as an explicit function of payoffs in Appendix A. The main result of the paper is summarized in the following theorem.

THEOREM 2 (Characterization of Dynamically Stable States): *Suppose that Assumptions* 1, 2, 4, *and* 5 *hold. Then there exists $\beta_0 \in (0,1)$ such that for all $\beta > \beta_0$, the following is true.*

(i) *For any $\phi \in \Phi$ there exists a set of protocols $\{\pi_s\}_{s \in \mathcal{S}}$ and a pure-strategy MPE σ of the game such that for any $s_0 \in \mathcal{S}$, $s_t^\sigma = \phi(s_0)$ for any $t \geq 1$; that is, the game reaches $\phi(s_0)$ after one period and stays in this state thereafter. Therefore, for each $s_0 \in \mathcal{S}$, $s = \phi(s_0)$ is a dynamically stable state.*

(ii) *Moreover, for any set of protocols $\{\pi_s\}_{s \in \mathcal{S}}$ there exists a pure-strategy MPE. Any such MPE σ has the property that there exists $\phi \in \Phi$ such that for any initial state $s_0 \in \mathcal{S}$, $s_t^\sigma = \phi(s_0)$ for all $t \geq 1$. Therefore, all dynamically stable states are axiomatically stable.*

(iii) *If, in addition, Assumption 3 holds, then the MPE is essentially unique: for any set of protocols $\{\pi_s\}_{s \in \mathcal{S}}$, any pure-strategy MPE σ, any initial state $s_0 \in \mathcal{S}$, and any $\phi \in \Phi$, $s_0^\sigma \sim \phi(s_0)$.*

Parts (i) and (ii) of Theorem 2 state that the set of dynamically stable states and the set of stable states $\mathcal{D}$ defined by axiomatic characterization in Theorem 1 coincide; any mapping $\phi \in \Phi$ that satisfies Axioms 1–3 is the outcome of a pure-strategy MPE and any such MPE implements the outcome of some $\phi \in \Phi$. An important implication is that the recursive characterization of axiomatically stable states in equation (5) can be used to calculate dynamically stable states.

The equivalence of the results of Theorems 1 and 2 is intuitive. Had players been short-sighted (impatient), they would care mostly about the payoffs in the next state or the next few states that would arise along the equilibrium path. When players are sufficiently patient, however ($\beta > \beta_0$), they care more about payoffs in the ultimate state than the payoffs along the transitional states. Consequently, winning coalitions are not willing to move to a state that is not (axiomatically) stable according to Theorem 1.

The proof of Theorem 2 is technically involved, but the idea is intuitive. For a given mapping $\phi \in \Phi$, we conjecture the continuation payoffs from accepting a particular alternative z in state s. We construct an MPE in the truncated game starting in state s in period t with terminal payoffs given by the continuation payoffs. We then show that transitions are given by ϕ, and the continuation payoffs are as conjectured. Conversely, if σ is an MPE, we show that transitions starting from any state s will eventually converge to some state $\psi(s)$, and then use Assumption 2(i) to show that any equilibrium path must lead to a state that is payoff-equivalent to $\psi(s)$. Finally, we verify that mapping $\psi(s)$ satisfies Axioms 1–3.

As illustrated by Example 1 in the Introduction, there is a tension between distribution of payoffs in a state and distribution of political power in the same state. Sometimes, Pareto improving transitions are impossible without changing the balance of political power. The next theorem clarifies the conditions under which Pareto efficiency will arise.

THEOREM 3 (Pareto Efficiency): *Suppose that for every two states x and y there is a state z such that $\{w_i(z)\}_{i \in \mathcal{I}} = \{w_i(y)\}_{i \in \mathcal{I}}$ and $\mathcal{W}_z \subset \mathcal{W}_x$, and no state is exogenously stable (i.e., $\mathcal{W}_s \neq \varnothing$ for each $s \in \mathcal{S}$). Then, every (axiomatically or dynamically) stable state is Pareto efficient. Otherwise, stable states may be Pareto inefficient.*

The positive result is that whenever the political environment is such that the current decision-makers can alter the economic allocation without giving up political power (which is captured here by the fact that a transition from x to z achieves the same payoffs as a transition to y without reallocating power to other groups), only Pareto efficient states are stable.

IV. Ordered States and Agents

Theorems 1 and 2 provide a complete characterization of axiomatically and dynamically stable states as a function of the initial state $s_0 \in \mathcal{S}$ provided that Assumptions 1 and 2 are satisfied. While the former is a very natural assumption and easy to check, Assumption 2 may be somewhat more difficult to verify. In this section, we show that when the sets of states $\mathcal{S}$ and agents $\mathcal{I}$ admit a linear order according to which individual stage payoffs satisfy single-crossing or single-peakedness properties (and the set of winning coalitions $\{\mathcal{W}_s\}_{s \in \mathcal{S}}$ satisfies some natural additional conditions), Assumption 2 is satisfied. This result enables more straightforward application of our main theorems in a wide variety of circumstances.

In a number of applications, the set of states $\mathcal{S}$ has a natural order, so that any two states x and y can be ranked. When such an order exists, we can take, without loss of any generality, $\mathcal{S}$ to be a subset of $\mathbb{R}$. Similarly, let $\mathcal{I} \subset \mathbb{R}$. Given these orders on the set of states and the set of individuals, we introduce certain standard restrictions on preferences.[20] All of the following restrictions and definitions refer to stage payoffs and are thus easy to verify.

DEFINITION 3 (Single-Crossing and Single-Peakedness): *Given $\mathcal{I} \subset \mathbb{R}$, $\mathcal{S} \subset \mathbb{R}$, and $\{w_i(s)\}_{i \in \mathcal{I}, s \in \mathcal{S}}$, the* **single-crossing condition** *holds if for any $i, j \in \mathcal{I}$ and $x, y \in \mathcal{S}$ such that $i < j$ and $x < y$, $w_i(y) > w_i(x)$ implies $w_j(y) > w_j(x)$ and $w_j(y) < w_j(x)$ implies $w_i(y) < w_i(x)$.*

Given $\mathcal{S} \subset \mathbb{R}$ and $\{w_i(s)\}_{i \in \mathcal{I}, s \in \mathcal{S}}$, preferences are **single-peaked** *if for any $i \in \mathcal{I}$ there exists a state x_i such that for any $y, z \in \mathcal{S}$, $y < z \leq x_i$ or $x_i \geq z > y$ implies $w_i(y) \leq w_i(z)$.*

We next introduce a generalization of the notion of the "median voter" to more general political institutions (e.g., those involving supermajority rules within the society or a club).

DEFINITION 4 (Quasi-Median Voter): *Given $\mathcal{I} \subset \mathbb{R}$ and $\{\mathcal{W}_s\}_{s \in \mathcal{S}}$, player $i \in \mathcal{I}$ is a* **quasi-median voter** *(in state s) if for any $X \in W_s$ such that $X = \{j \in \mathcal{I} : a \leq j \leq b\}$ for some $a, b \in \mathbb{R}$ we have $i \in X$.*

Denote the set of quasi-median voters in state s by M_s. Lemma 1 in the proof of Theorem 4 shows that, provided that Assumption 1 is satisfied, this set is nonempty.

[20] Rothstein (1990) and Austen-Smith and Banks (1999) study another restriction, order-restricted preferences. As Gans and Smart (1996) show, this notion is equivalent to single-crossing and is thus covered by our framework.

DEFINITION 5 (Monotonic Median Voter Property): *Given* $\mathcal{I} \subset \mathbb{R}$ *and* $\mathcal{S} \subset \mathbb{R}$, *the set of winning coalitions* $\{\mathcal{W}_s\}_{s\in\mathcal{S}}$ *has* **monotonic median voter property** *if for each* $x, y \in \mathcal{S}$ *satisfying* $x < y$ *there exist* $i \in M_x, j \in M_y$ *such that* $i \leq j$.

The last definition is general enough to encompass majority and supermajority voting as well as those voting rules that apply to a subset of players (such as club members or those that are part of a restricted franchise). Finally, we also impose the following weak genericity assumption.

ASSUMPTION 6 (Weak Genericity): *Preferences* $\{w_i(s)\}_{i\in\mathcal{I},s\in\mathcal{S}}$ *and the set of winning coalitions* $\{\mathcal{W}_s\}_{s\in\mathcal{S}}$ *are such that for any* $x, y, z \in \mathcal{S}$, $x \succeq_z y$ *implies* $x \succ_z y$ *or* $x \sim y$.

Assumption 6 is satisfied if no player is indifferent between any two states (though it does not rule out such indifferences). Next, we present the main result of this section.

THEOREM 4 (Characterization with Ordered States): *For any* $\mathcal{I} \subset \mathbb{R}$, $\mathcal{S} \subset \mathbb{R}$, *preferences* $\{w_i(s)\}_{i\in\mathcal{I},s\in\mathcal{S}}$, *and winning coalitions* $\{\mathcal{W}_s\}_{s\in\mathcal{S}}$ *satisfying Assumption* 1 *and Assumption* 6*:*

(i) *If single-crossing condition and monotonic median voter property hold, then Assumption* 2 *is satisfied and, thus, Theorems* 1 *and* 2 *apply.*

(ii) *If preferences are single-peaked and for any* $x, y \in \mathcal{S}$ *and any* $X \in \mathcal{W}_x$, $Y \in \mathcal{W}_y$ *we have* $X \cap Y \neq \varnothing$, *then Assumption* 2 *is satisfied and thus Theorems* 1 *and* 2 *apply.*

Part (ii) of Theorem 4 requires a stronger condition than the monotonic median voter property. Because this condition implies the monotonic median voter property, part 1 of the theorem continues to be true under the hypothesis of part 2. The converse is not true, however.

V. Application

In this section, we apply our results to the dynamics of political rights discussed in Example 2 in the Introduction. Consider a society $\mathcal{I} = \{1, \ldots, n\}$ consisting of n groups (or individuals) ranked in ascending order of religiosity, so that 1 is most secular and n is most religious. There is a one-dimensional policy space indexed by $\rho \in R = \{\rho^1, \ldots, \rho^r\}$, where higher ρ corresponds to greater tolerance towards religiosity and less tolerance towards nonreligious individuals.

In each period t, the set of individuals who have the right to political participation is Z_t, a *connected* subset of $\mathcal{I}$. We assume that at each date, political decisions are made by α-(super)majorities (i.e., coalitions of at least $\alpha|Z_t|$ members). These decisions include the determination of which subset of the society will have the right for political participation in the next period (i.e., the subset Z_{t+1}) and the next period's religiosity policy ρ_{t+1}. The *state* can thus be represented by $s = (\rho, Z)$ where $\rho \in R$ and Z is a connected subset of $\mathcal{I}$.

We assume that each individual cares about the policy toward religiosity ρ and also about the extent of political participation in society. For example, higher political participation may increase income or the amount of public goods, or decrease political instability. Since these effects are likely to affect all players equally, we assume that preferences over states are given by

$$w_i(s) = v_i(\rho) + V(Z),$$

where $V(Z)$ is any function, and $v_i(\rho)$ satisfies the strict increasing differences condition:

$$v_i(\rho) - v_j(\rho) \text{ is strictly increasing in } \rho \text{ whenever } i > j.$$

This condition implies, in particular, that the sequence of ideal policies of agents, $\{\hat{\rho}_i\}_{i=1}^n$, is (weakly) monotonically increasing. It is satisfied, for example, when individuals have quadratic utility function $v_i(\rho) = -(\rho - \hat{\rho}_i)^2$.

Since an α-(super)majority in Z chooses the religiosity policy for the next period, ρ, it is natural that this policy choice is between $\hat{\rho}_{\min M_Z}$ and $\hat{\rho}_{\max M_Z}$, where, as before, M_Z is the set of quasi-median voters. Formally, the set of states $\mathcal{S}$ consists of all pairs (ρ, Z), where Z is a connected subset of $\mathcal{I}$ and $\hat{\rho}_{\min M_Z} \leq \rho \leq \hat{\rho}_{\max M_Z}$.

This example specifies a rich and highly complex social situation. Granting political participation to previously excluded religious (resp., secular) individuals will have short-run economic benefits, but could unleash a political process that might later on deprive secular (resp., religious) individuals of their political rights. The richness of the environment results from the fact that individuals with political rights are simultaneously choosing a policy ρ and the subset of the society Z that will have political rights in the future. Despite this, the tools and insights developed so far can be applied to derive a sharp characterization of the structure of equilibria.

We first establish that Assumptions 1 and 2 are satisfied, so that the dynamic equilibrium in this environment can be characterized by applying Theorems 1 and 2. To simplify the exposition of the results, we assume that $w_i(s) \neq w_i(s')$ for any $i \in \mathcal{I}$ and $s \neq s'$, which ensures that Assumption 6 holds. Thus, we can use $\phi(s_0)$ to denote the state, both axiomatically and dynamically stable, that corresponds to intitial state s_0.

PROPOSITION 1:

(i) *For any degree of (super) majority* α, *Assumptions* 1 *and* 2 *are satisfied and thus Theorems* 1 *and* 2 *apply in this environment. In particular, there exists* $\beta_0 < 1$ *such that for any discount factor* $\beta > \beta_0$, *an equilibrium exists.*

(ii) *Assume* $V(Z)$ *to be (strictly) increasing (whenever* $Z \neq Z'$, $Z \subset Z'$ *implies* $V(Z) < V(Z')$*). Then for any initial state* s_0, $\phi(s_0) = s = (Z, \rho)$ *with* Z *containing at least one of the extreme players,* 1 *or* n.

To prove the first part, we enumerate the states $\{s_1, \ldots, s_{|\mathcal{S}|}\}$ so that ρ_s is weakly increasing. We then establish that strict increasing differences and monotonic median voter properties hold, and use Theorem 4.[21] The second part of Proposition 1 shows that when V is an increasing function, stable states provide political rights to at least one of the extreme members of the society. Intuitively, this holds because the threat to the current set of individuals holding power comes either from greater religiosity or greater secularism. Thus, there will necessarily be expansion toward the less threatening side.

This result does not rule out that political rights will be given to everybody in society. The next proposition studies this question. In what follows, we assume that $V(Z)$ is strictly increasing and $v_i(\rho)$ is single-peaked for all $i \in \mathcal{I}$.

PROPOSITION 2: *Define* $A \equiv V(\mathcal{I}) - \max_{i \in \mathcal{I}} V(\mathcal{I} \backslash \{i\})$ *and* $A_i \equiv V(\mathcal{I}) - V(\{i\})$.

(i) *Suppose* $v_1(\hat{\rho}_1) - v_1(\hat{\rho}_{\min M_{\mathcal{I}}}) < A$ *and* $v_n(\hat{\rho}_n) - v_n(\hat{\rho}_{\max M_{\mathcal{I}}}) < A$. *Then for any initial state* s, $Z(\phi(s)) = \mathcal{I}$.

(ii) *Suppose* $v_1(\hat{\rho}_1) - v_1(\hat{\rho}_{\min M_{\mathcal{I}}}) > A_1$ *and* $v_n(\hat{\rho}_n) - v_n(\hat{\rho}_{\max M_{\mathcal{I}}}) > A_n$. *There exists* $k \in \mathbb{N}$ *such that if the initial state* s_0 *satisfies* $|Z(s_0)| \leq k$, *then* (i) *when* Z_0 *includes the middle player* (*or at least one of the two middle players if* n *is even*), $Z(\phi(s_0)) = \mathcal{I}$, *and* (ii) *when* Z_0 *includes one of the extreme players*, $Z(\phi(s_0)) \neq \mathcal{I}$.

(iii) *If* $\alpha > \frac{n-1}{n}$, *i.e., the rule is unanimity, then for any initial state* s_0, $Z(\phi(s_0)) = \mathcal{I}$.

The first part of this proposition shows that if utility gains from greater political participation are sufficiently large (sufficient to compensate extremists for a change in policies towards religiosity), then political participation is granted to all parties. More interestingly, the second part shows that when these gains are not sufficiently large, political participation is granted to all if political power initially rests with moderates and is not granted if it rests with one of the extremes.

The third part asserts that if the decision rule is unanimity, then political rights can be extended to all individuals because the status quo religious policy may be preserved in this case. Intuitively, unanimity guarantees that political power will not shift to extremists of the opposite conviction and thus enables expansion of political participation.[22] This final result raises the question of whether the groups that are currently powerful can introduce a unanimity clause into the current constitution or set of rules in order to cement their political power even as reforms are implemented. While this may be feasible under certain circumstances, we believe that it is in general not possible to grant political participation to new groups and individuals but effectively take away their ability to implement significant future policy

[21] Note, however, that the original environment is not ordered and this theorem could not have been applied directly; we can only apply it after undertaking this enumeration.

[22] This is similar to the general result in Theorem 3, which shows that when a policy may be changed without undermining the power of currently powerful players, equilibria are necessarily Pareto efficient. (Recall that only states with full participation are Pareto efficient in this application).

changes by introducing unanimity clauses or other restrictions. (One reason is that this would go against the spirit of current allocation of political power determining current policy choices and reforms.)

Finally, we consider an even richer environment where individuals also choose the degree of (super)majority rule α. In particular, now a state is $s = (\rho, Z, \alpha)$, with $\alpha \in \mathcal{A}$, where $\mathcal{A} \subset [1/2, 1]$ is a finite set. Then, since $\mathcal{A}$ is a finite set, our previous results yield the following proposition.

PROPOSITION 3:

(i) *In this environment, Assumptions* 1 *and* 2 *are satisfied and thus Theorems* 1 *and* 2 *apply.*

(ii) *Suppose that* $\mathcal{A}$ *contains* $\alpha > \frac{n-1}{n}$. *Then for any equilibrium and for any state* s_0, $Z(\phi(s_0)) = \mathcal{I}$.

This proposition demonstrates the applicability of our results in the environment in which the degree of supermajority necessary for future decisions is also a collective choice. Again, whenever unanimity can be imposed, full participation is guaranteed. Intuitively, they can make any (and every) individual a veto player, preventing future policy changes. We should, however, reiterate at this point that this result does not imply that changing the decision rule to unanimity is always or often feasible. In many relevant situations, including those mentioned in the Introduction and several we discuss in the online Appendix, it is a hard-wired feature that future decisions will be made by a (weighted) majority of those who participate in future decision making, and their ability to change policies and laws cannot be restricted by past unanimity clauses or constitutional requirements.[23] Along these lines, for example, those worried about the "slippery slope" of giving more rights to religious groups in Turkey fear that any constitutional guarantees can be changed in the future.

VI. Conclusion

A central feature of collective decision making in many social situations, such as societies choosing their constitutions or institutions, leaders building political coalitions, countries joining international unions, or private clubs deciding on their membership, is that the rules that govern regulations and procedures for future decision making, and inclusion and exclusion of members, are made by the current members and under the current regulations. This feature implies that dynamic collective decisions must recognize the impact of current decisions on future choices.

We developed a framework for a systematic study of this class of problems. We provided both an axiomatic and a noncooperative characterization of stable states and showed that the set of (dynamically) stable states can be computed recursively. This recursive characterization highlights that a particular state s is stable if no other stable state makes a winning coalition (in s) better off. This implies that stable states

[23] See Acemoglu, Egorov, and Sonin (2008) for an example.

need not be Pareto efficient; there may exist a state that provides higher payoffs to all individuals, but is itself not stable.

Our analysis relies on several substantive and technical assumptions. Substantive assumptions, such as a minimum amount of acyclicity, are essential for our approach. Others, the technical ones, are adopted for convenience and can be relaxed, though often at the cost of further complication. Among possible extensions, most interesting might be to introduce stochastic elements so that the set of feasible transitions or the distribution of powers vary stochastically over time, and to include capital-like state variables so that some subcomponents of the state have autonomous dynamics.

Appendix

PROOF OF THEOREM 1:

Part 1: We first construct, by induction, a sequence of states $\{\mu_1, \ldots, \mu_{|\mathcal{S}|}\}$ such that

(A1) $$\text{if } 1 \leq j < l \leq |\mathcal{S}|, \text{ then } \mu_l \not\succ_{\mu_j} \mu_j.$$

Assumption 2(i) implies that for any nonempty collection of states $\mathcal{Q} \subset \mathcal{S}$, there exists $z \in \mathcal{Q}$ such that for any $x \in \mathcal{Q}$, $x \not\succ_z z$. Applying this result to $\mathcal{S}$, we obtain μ_1. Now, suppose we have defined μ_j for all $j \leq k - 1$, where $k \leq |\mathcal{S}|$. Applying the same result to the collection of states $\mathcal{S} \backslash \{\mu_1, \ldots, \mu_{k-1}\}$, we conclude that there exists μ_k satisfying (A1) for each k.

The second step is to construct, again by induction, a candidate mapping $\phi : \mathcal{S} \rightarrow \mathcal{S}$. For $k = 1$, let $\phi(\mu_k) = \mu_k$. Suppose we have defined $\phi(\mu_j)$ for all $j \leq k - 1$ where $2 \leq k \leq |\mathcal{S}|$. Define the collection of states $\mathcal{M}_k$ as in (4). This is the subset of states for which ϕ has already been defined and which satisfies $\phi(s) = s$ and is preferred to μ_k within μ_k. If $\mathcal{M}_k$ is empty, then we define $\phi(\mu_k) = \mu_k$. If $\mathcal{M}_k$ is nonempty, then take $\phi(\mu_k) = z \in \mathcal{M}_k$ such that

(A2) $$s \not\succ_{\mu_k} z \text{ for any } s \in \mathcal{M}_k$$

(applying Assumption 2(ii) to $\mathcal{M}_k$, we get that there exists $z \in \mathcal{M}_k$ such that $s \not\succeq_{\mu_k} z$, and thus $s \not\succ_{\mu_k} z$, for all $s \in \mathcal{M}_k$). Proceeding inductively for all $2 \leq k \leq |\mathcal{S}|$, we obtain ϕ as in equation (5).

To complete the proof, we need to verify that mapping ϕ in (5) satisfies Axioms 1–3. This is straightforward for Axioms 1 and 2. In particular, by construction, either $\phi(\mu_k) = \mu_k$ (in that case these axioms hold trivially), or $\phi(\mu_k)$ is an element of $\mathcal{M}_k$. In the latter case, $\phi(\mu_k) \succ_{\mu_k} \mu_k$ and $\phi(\phi(\mu_k)) = \phi(\mu_k)$ by (4). To check Axiom 3, suppose that for some state μ_k there exists y such that $y \succ_{\mu_k} \mu_k$, $y = \phi(z)$, and $y \succ_{\mu_k} \phi(\mu_k)$. Then $y \succ_{\mu_k} \mu_k$, combined with condition (A1), implies that $y \in \{\mu_1, \ldots, \mu_{k-1}\}$, and therefore $y \in \mathcal{M}_k$. But then $y \succ_{\mu_k} \phi(\mu_k)$ contradicts (A2). This means that such y does not exist, and therefore Axiom 3 is satisfied.

Part 2: This statement is equivalent to the following: if, given a sequence $\{\mu_1, \ldots, \mu_{|\mathcal{S}|}\}$ with the property (A1), $\phi(\mu_k)$ does not satisfy (5) for some k, then ϕ

does not satisfy Axioms 1–3. Suppose first that $\phi(\mu_k)$ is not given by (5) at $k = 1$. Then $\phi(\mu_1) \neq \mu_1$, so $\phi(\mu_1) = \mu_l$ for $l > 1$. In this case, ϕ does not satisfy Axiom 1, because $\mu_l \not\succ_{\mu_1} \mu_1$ by (A1). Now, let $k > 1$ be the smallest k for which $\phi(\mu_k)$ is not given by (5). Suppose, to obtain a contradiction, that Axioms 1–3 hold. Then $\mathcal{M}_k$ in (4) is well defined, and either $\mathcal{M}_k = \varnothing$ or $\mathcal{M}_k \neq \varnothing$. If $\mathcal{M}_k = \varnothing$ and $\phi(\mu_k)$ is not given by (5), then $\phi(\mu_k) \neq \mu_k$. Then, Axioms 1 and 2 imply $\phi(\mu_k) \succ_{\mu_k} \mu_k$ and $\phi(\phi(\mu_k)) = \phi(\mu_k)$. Since $\mathcal{M}_k = \varnothing$, we must have that $\phi(\mu_k) = \mu_l$ for $l > k$, but in this case $\phi(\mu_k) \succ_{\mu_k} \mu_k$ contradicts (A1). This contradiction implies that ϕ violates either Axiom 1 or Axiom 2 (or both). If $\mathcal{M}_k \neq \varnothing$, then consider $\mu_l = \phi(\mu_k)$. If $l > k$, then Axiom 1 is violated. If $l = k$, then ϕ violates Axiom 3 (to see this, take any $z \in \mathcal{M}_k \neq \varnothing$ and observe that $z \succ_{\mu_k} \mu_k$, $z \succ_{\mu_k} \phi(\mu_k)$ and $\phi(z) = z$). If $l < k$, then Axiom 1 and Axiom 2 imply $\phi(\mu_k) \in \mathcal{M}_k$. Then, since $\phi(\mu_k)$ is not given by (5), there exists some $y \in \mathcal{M}_k$ such that $y \succ_{\mu_k} \phi(\mu_k)$. But in this case ϕ violates Axiom 3, since $y \succ_{\mu_k} \phi(\mu_k)$, $y \succ_{\mu_k} \mu_k$, and $\phi(y) = y$. We have obtained contradictions in all possible cases.

Part 3: Suppose, to obtain a contradiction, that $\mathcal{D}_{\phi_1} \neq \mathcal{D}_{\phi_2}$. Then $\exists k : 1 \leq k \leq |\mathcal{S}|$ such that $\mu_j \in \mathcal{D}_{\phi_1} \Leftrightarrow \mu_j \in \mathcal{D}_{\phi_2}$ for all $j < k$, but either $\mu_k \in \mathcal{D}_{\phi_1}$ and $\mu_k \notin \mathcal{D}_{\phi_2}$, or $\mu_k \notin \mathcal{D}_{\phi_1}$ and $\mu_k \in \mathcal{D}_{\phi_2}$. Without loss of generality, assume that $\mu_k \in \mathcal{D}_{\phi_1}$ and $\mu_k \notin \mathcal{D}_{\phi_2}$. Then part 2 implies that $\phi_2(\mu_k) = \mu_l$ for some $l < k$. Applying Axioms 1 and 2 to mapping ϕ_2, we obtain $\mu_l \succ_{\mu_k} \mu_k$ and $\phi_2(\mu_l) = \mu_l$; the latter implies that $\mu_l \in \mathcal{D}_{\phi_2}$. Since, by hypothesis, $\mu_j \in \mathcal{D}_{\phi_1} \Leftrightarrow \mu_j \in \mathcal{D}_{\phi_2}$ for all $j < k$, we have $\mu_l \in \mathcal{D}_{\phi_1}$. Therefore, $\mu_l \succ_{\mu_k} \mu_k$, $\mu_l \succ_{\mu_k} \phi_1(\mu_k)$ (because $\phi_1(\mu_k) = \mu_k$), and $\phi_1(\mu_l) = \mu_l$, but this violates Axiom 3 for mapping ϕ_1.

Part 4: Suppose Assumption 3 holds. Suppose, to obtain a contradiction, that for some state s, $\phi_1(s) \nsim \phi_2(s)$. Part 3 of this Theorem implies that $\phi_1(s) = s \Leftrightarrow \phi_2(s) = s$; since $\phi_1(s) \nsim \phi_2(s)$, we must have that $\phi_1(s) \neq s \neq \phi_2(s)$. Axiom 1 then implies $\phi_1(s) \succ_s s$, $\phi_2(s) \succ_s s$, and Assumption 3 implies that either $\phi_1(s) \succ_s \phi_2(s)$ or $\phi_2(s) \succ_s \phi_1(s)$. Without loss of generality, suppose that the former is the case. Then for $y = \phi_2(s)$ there exists $z = \phi_1(s)$ such that $z \succ_s y$, $z \succ_s s$, and $\phi_2(z) = z$ (the latter holds because $\phi_1(s) = s$ by Axiom 2, and then $\phi_2(s) = s$ by part 3 of this Theorem). Then we can apply Axiom 3 to ϕ_2 and s and conclude that $\phi_2(s) \neq y$, a contradiction.

PROOF OF THEOREM 2:

Part 1: Assume β satisfies the following conditions:

(A3) $$\text{for any } i \in \mathcal{I} \quad \text{and} \quad x, y \in \mathcal{S},$$

$$w_i(x) < w_i(y) \quad \text{implies } w_i(x) < \left(1 - \beta^{|\mathcal{S}|}\right)\tilde{w}_i + \beta^{|\mathcal{S}|} w_i(y).$$

To prove part 2, we will also need the following conditions:

(A4) $$\text{for any } i \in \mathcal{I} \quad \text{and} \quad x, y, z \in \mathcal{S},$$

$$w_i(x) < w_i(y) \quad \text{implies } \frac{1 - \beta}{\beta}\left(w_i(z) - (1 - \beta)\tilde{w}_i\right) + \beta w_i(x) < w_i(y).$$

In total, there is a finite number of conditions in (A3) and (A4). Therefore, there exists $\beta_0 \in (0,1)$ such that for all $\beta > \beta_0$, (A3) and (A4) hold.

Pick any $\phi \in \Phi$ and any $s_0 \in S$. We construct an MPE of the game such that for each period $t \geq 1$, $s_t = \phi(s_{t-1})$. For $i \in \mathcal{I}$ and $s, q \in \mathcal{S}$, let

$$\text{(A5)} \qquad V_i(s,q) = \begin{Bmatrix} (1-\beta)\, w_i(s) & \text{if } s = q \\ (1-\beta)\, \tilde{w}_i & \text{if } s \neq q \end{Bmatrix} + \begin{Bmatrix} \beta w_i(\phi(q)) & \text{if } \phi(q) = q \\ \beta(1-\beta)\,\tilde{w}_i + \beta^2 w_i(\phi(q)) & \text{if } \phi(q) \neq q \end{Bmatrix}.$$

In the equilibrium we construct below, $V_i(s,q)$ will be the continuation payoff of i as a function of the current state s and the accepted proposal q. In the remainder, we drop time indices.

For each $s \in \mathcal{S}$, take $K_s \geq |\mathcal{S}| - 1$. Take $\pi_s(\cdot)$ such that Assumption 5 holds, and if $\phi(s) \neq s$, then $\pi_s(K_s) = \phi(s)$. Consider strategy profile σ^* constructed as follows: Each $i \in \mathcal{I}$ votes for proposal P_k (says *yes*) if and only if:

(i) either $k = K_s$ (we are at the last stage of voting), $P_{K_s} = \phi(s)$ and $V_i(s,\phi(s)) > V_i(s,s)$;

(ii) or $V_i(s,P_k) > V_i(s,\phi(s))$.

In addition, if $\pi_s(k) \in \mathcal{I}$ for some k, this player chooses proposal P_k arbitrarily.

The strategy profile σ^* is Markovian. We will show that it is an MPE in three steps.

First, we show that under the strategy profile σ^*, there is a transition to $\phi(s)$ if $\phi(s) \neq s$ and no transition if $\phi(s) = s$. Suppose that $\phi(s) \neq s$, then Axiom 1 implies that

$$X_s \equiv \{i : w_i(\phi(s)) > w_i(s)\} \in \mathcal{W}_s.$$

Now, (A3) and $\beta > \beta_0$ imply that for all $i \in X_s$, we have

$$V_i(s,\phi(s)) = (1-\beta)\,\tilde{w}_i + \beta w_i(\phi(s)) > (1-\beta)\, w_i(s) + \beta(1-\beta)\,\tilde{w}_i + \beta^2 w_i(\phi(s)) = V_i(s,s).$$

Consequently, if $\phi(s) \neq s$, then under σ^*, there is transition to $\phi(s)$ if stage K_s is reached.

Let us now show that there exist no $X'_s \in \mathcal{W}_s$ and $P_k \in \mathcal{S}$ such that $V_i(s,P_k) > V_i(s,\phi(s))$ for all $i \in X'_s$; i.e., the set of players for whom $V_i(s,P_k) > V_i(s,\phi(s))$ is not a winning coalition in s. To obtain a contradiction, suppose there exists such a X'_s and P_k. Then, since $P_k \neq s$ and $\phi(\phi(s)) = \phi(s)$, we would have that for all $i \in X'_s$,

$$w_i(\phi(P_k)) > (1-\beta)\,\tilde{w}_i + \beta w_i(\phi(P_k)) \geq V_i(s,P_k) > V_i(s,\phi(s)) \geq (1-\beta)\,\tilde{w}_i + \beta w_i(\phi(s)),$$

and thus, by (A3),

$$w_i\big(\phi(P_k)\big) > w_i\big(\phi(s)\big) \quad \text{for all } i \in X'_s.$$

So, $X'_s \in \mathcal{W}_s$ implies $\phi(P_k) \succ_s \phi(s)$, which, given that $\phi(s) \succ_s s$, yields $\phi(P_k) \succ_s s$ by Assumption 2(*ii*). But $\phi(P_k) \succ_s \phi(s)$, $\phi(P_k) \succ_s s$, and $\phi(\phi(P_k)) = \phi(P_k)$ contradicts Axiom 3. Therefore, the set of players with $V_i(s, P_k) > V_i\big(s, \phi(s)\big)$ does not form a winning coalition in s. This means that under σ^*, no proposal is accepted if $\phi(s) = s$, and if $\phi(s) \neq s$, then no proposal is accepted in all stages but the last one, and in the last stage $P_{K_s} = \phi(s)$ is accepted.

Second, we verify that given σ^*, continuation payoffs after acceptance of proposal q are given by (A5). If proposal $q \neq s$ is accepted, then there is an immediate transition to q, while if no proposal is accepted, then each player i receives stage utility $(1 - \beta)\, w_i(s)$. In the next period, there is a transition $\big(\text{to } \phi(q)\big)$ under σ^* if and only if $\phi(q) \neq q$, and after that there are no transitions along the equilibrium path. Hence, the continuation payoffs are given by (A5).

Third, we show that there are no profitable deviations from σ^* at any stage. For an agenda-setter, this holds because no proposal that he can make is accepted. For a voter, notice that since continuation strategies are Markovian, it is always a best response to vote for the option that the player (weakly) prefers, and this is what profile σ^* prescribes. Indeed, if $\phi(s) \neq s$, then in the last voting stage, each player i compares continuation payoff $V_i\big(s, \phi(s)\big)$ if the proposal is accepted and $V_i(s,s)$ if it is rejected. In all other voting stages, player i receives $V_i(s, P_k)$ if proposal P_k is accepted and $V_i\big(s, \phi(s)\big)$ if it is rejected (because $\phi(s)$ will eventually be accepted if $\phi(s) \neq s$ and no proposal is accepted if $\phi(s) = s$). Therefore, there are no profitable deviations from σ^* given the continuation payoffs in (A1). Thus, σ^* is a best response to itself at every voting stage for any $s \in \mathcal{S}$, and thus σ^* is an MPE of the entire game.

Part 2: We first prove that an MPE exists, and then that any MPE has the stated properties. We first construct a mapping ϕ satisfying Axioms 1–3. Take a sequence of states $\{\mu_1, \ldots, \mu_{|\mathcal{S}|}\}$ satisfying (A1). Then, follow the procedure described in Theorem 1. First, we set $\phi(\mu_1) = \mu_1$. If for $l \geq 2$ we have $\mathcal{M}_l = \varnothing$, then $\phi(\mu_l) = \mu_l$; otherwise, define $Z_l \subset \mathcal{M}_l$ by

$$Z_l = \{z \in \mathcal{M}_l : \forall s \in \mathcal{M}_l : s \nsim z \Rightarrow s \not\succ_{\mu_l} z\}.$$

Then $Z_l \neq \varnothing$, as we can apply Assumption 2(b) to $\mathcal{M}_l$. Choose a particular element of Z_l as $\phi(\mu_l)$ as follows. Let Y_{μ_l} be the set of stages of protocol π_{μ_l} such that for any stage $j \in Y_{\mu_l}$, $\pi_{\mu_l}(j) \in \mathcal{S}$ implies $\pi_{\mu_l}(j) \in Z_l$, and $\pi_{\mu_l} \in \mathcal{I}$ implies that for some $z \in Z_l : w_i(z) > w_i(\mu_l)$, where $i = \pi_{\mu_l}(j)$. By Assumption 5, Y_{μ_l} is nonempty; let $k^*_{\mu_l}$ be the last stage from Y_{μ_l}. If $\pi_{\mu_l}(k^*_{\mu_l}) \in \mathcal{S}$, then let $\phi(\mu_l) = \pi_{\mu_l}(k^*_{\mu_l})$, while if $\pi_{\mu_l}(k^*_{\mu_l}) \in \mathcal{I}$, then let $\phi(\mu_l)$ be any element $z \in Z_l$ such that $w_i(z) > w_i(\mu_l)$ for $i = \pi_{\mu_l}(j)$. Proceeding inductively by l, we get mapping ϕ.

We now construct an equilibrium that implements ϕ. In this equilibrium, continuation payoff of player i if the current state is s and proposal q is accepted, $V_i(s,q)$, is given by (A5); and if no alternative is accepted, each player i receives $V_i(s,s)$. Given these continuation payoffs, each period can be viewed as a finite (truncated) game with terminal payoffs given by $V_i(s,q)$. We construct an MPE σ' of this truncated game by backward induction.

Case (i): $\phi(s) \neq s$. Given any current states s, consider the stage k_s^* defined above in the construction of mapping ϕ. If k_s^* is not the last stage, then for stages from K_s down to $k_s^* + 1$ we do the following. Suppose that in the last stage, the voting is over the alternative s'. Comparing payoffs as in the proof of part 1, we see that in an SPE, s' must be accepted if and only if $s' = \phi(s)$ and rejected otherwise. But by definition of Y_s, s' may be voted only if nominated by some player i. Proceeding backward to the agenda-setting stage, we notice that such player i must have $w_i(\phi(s)) \leq w_i(s)$, and then he strictly prefers to stay in s, which means that nominating $s' = \phi(s)$ is not his best action. By not nominating $\phi(s)$ if the game reached the last stage K_s he ensures that the next state is s. We can apply the same reasoning to all voting stages up to $k_s^* + 1$, and get an SPE in the subgame starting from stage $k_s^* + 1$ where no proposal is accepted and s is implemented.

Consider now stage k_s^*. By the same reasoning, only $\phi(s)$ may be accepted if nominated. At this stage, it either happens automatically according to the protocol or, if $\pi_s(k_s^*) = i \in \mathcal{I}$, then i's best response is to nominate $\phi(s)$: if i does not, then s persists for an extra period. Hence, in a subgame that starts at stage k_s^*, there is a SPE where $\phi(s)$ is accepted.

If $k_s^* \neq 1$, we proceed with backward induction. At stage $k_s^* - 1$, no proposal other than $\phi(s)$ may be accepted, and we can choose voting strategies such that $\phi(s)$ is rejected at this stage (it is later accepted at stage k_s^*). If at this stage the agenda-setter is some player i, he is indifferent, and we pick any action. Proceeding backward, we finish constructing an SPE σ' of this truncated game if the current state is s for the case $\phi(s) \neq s$.

Case (ii): $\phi(s) = s$. Take the last voting stage, and suppose that some proposal $s' \neq s$ is considered. For a player i to vote for s', $w_i(\phi(s')) > w_i(s)$ must hold. Since $\phi(s) = s$, however, such players do not form a winning coalition. Consequently, we can choose voting strategies so that a transition to another state will not be supported. Consequently, at the agenda-setting stage, any action may be chosen, as none of his proposals may be accepted. We can use backward induction to construct a strategy profile σ' where no proposal is accepted.

Note that in both cases, we can choose σ' to be Markovian by choosing the same actions in equivalent subgames for any player who is indifferent. Having done so for all $s \in \mathcal{S}$, we get a Markovian strategy profile σ. But given that in this strategy profile all transitions are one-stage, the payoffs are indeed given by (A5), and therefore there is no profitable one-shot deviation (otherwise, σ' would not be an SPE for some s). This shows that σ is an MPE.

Our next step is to establish the properties that any MPE satisfies. Take any set of protocols $\{\pi_s(\cdot)\}_{s\in\mathcal{S}}$ and any pure-strategy MPE σ. For any state s, the proposal q that is accepted along the equilibrium path is well-defined (let $q = s$ if all proposals are rejected) and define $\chi(s) = q$. First, note that $\chi : \mathcal{S} \to \mathcal{S}$ has "no cycles": if $\chi(s) \neq s$ then for any $n > 1$, $\chi^n(s) \neq s$ (where $\chi^2(s) \equiv \chi(\chi(s))$, etc.). This can be established by contradiction. Suppose there exists n such that $\chi^n(s) = s$, but

$\chi(s) \neq s$. Denote by $J_s \subset \{1, \ldots, K_s\}$ the set of voting stages in state s where a proposal P_k made along the equilibrium path is accepted. By definition of χ, the first voting stage in J_s leads to $\chi(s)$. Two cases are possible.

Case (i): for every $k \in J_s$, $\chi^{n+1}(P_k) \neq \chi^n(P_k)$ for all n. Then consider the last voting stage $k' \in J_s$. If $P_{k'}$ is accepted, each player i receives $\tilde{w}_i$, and if $P_{k'}$ is rejected, i gets $(1 - \beta) w_i(s) + \beta \tilde{w}_i > \tilde{w}_i$. But $P_{k'}$ cannot be accepted in an MPE, yielding the desired contradiction.

Case (ii): for some $k \in J_s$, $\chi^{n+1}(P_k) = \chi^n(P_k)$ for some n. Denote the set of such k by $J_s' \subset J_s$; clearly, the first stage in J_s is not in J_s'. Let k' be the first stage in J_s'; then $\chi^{n+1}(P_{k'}) = \chi^n(P_{k'})$ for all $n \geq |\mathcal{S}| - 1$. Consider the stage k'' in J_s that precedes k'. Accepting the proposal made at k'', $P_{k''}$, gives $\tilde{w}_i$ to each player i, while rejecting it yields at least $(1 - \beta^{|\mathcal{S}|}) \tilde{w}_i + \beta^{|\mathcal{S}|} w_i(\chi^{|\mathcal{S}|}(P_{k''})) > \tilde{w}_i$. Therefore, proposal $P_{k''}$ cannot be accepted in any MPE, which yields a contradiction and establishes the "no cycle" result.

This "no cycle" result in turn implies that $\chi^n(s) = \chi^{|\mathcal{S}|-1}(s)$ for all $n \geq |\mathcal{S}| - 1$. Define $\psi(s) = \chi^{|\mathcal{S}|-1}(s)$, and, with the convention that $\chi^0(s) \equiv s$,

$$m(s) = \min\{n \in \mathbb{N} \cup \{0\} : \chi^n(s) = \psi(s)\}. \tag{A6}$$

Evidently, $0 \leq m(s) \leq |\mathcal{S}| - 1$, and $m(s) = 0$ if and only if $\psi(s) = \chi(s) = s$. Moreover,

$$\psi(\psi(s)) = \chi(\psi(s)) = \psi(\chi(s)) = \psi(s) \tag{A7}$$

for any state s, as follows from the definition of mapping ψ. Finally, define

$$\overline{V}_i(s) = \left\{\begin{array}{ll}(1 - \beta) w_i(s) & \text{if } \chi(s) = s \\ (1 - \beta) \tilde{w}_i & \text{if } \chi(s) \neq s\end{array}\right\} + \beta w_i(\chi(s)), \tag{A8}$$

which is the equilibrium payment of player i if the equilibrium proposal $\chi(s)$ is accepted, and, slightly abusing the notation, $\overline{V}_i$,

$$\overline{V}_i(s,q) = \left\{\begin{array}{ll}(1 - \beta) w_i(s) & \text{if } s = q \\ (1 - \beta) \tilde{w}_i & \text{if } s \neq q\end{array}\right\} + \beta \overline{V}_i(q). \tag{A9}$$

Clearly, $\overline{V}_i(s,q)$ gives the continuation payoff of player i if in state s alternative q is accepted, and equilibrium play (according to σ) follows. We now prove an auxiliary result; then we will prove that $\psi(s)$ satisfies Axioms 1 and 2, then that $\chi(s) = \psi(s)$ (which implies $s_t = \chi(s_0)$ for all $t \geq 1$), and finally that ψ satisfies Axiom 3.

Proof that if proposals P_{k_j} and P_{k_l}, $j < l$, are proposed and accepted in state s, then $\psi(P_{k_j}) \sim \psi(P_{k_l})$ and $m(P_{k_j}) \leq m(P_{k_l})$. We only need to consider the case where $\chi(s) \neq s$, and thus, $m(s) \geq 1$. For each state s take the set of voting stages J such that for each $k \in J$, the proposal P_k is accepted. Let $J = \{k_1, \ldots, k_{|J|}\}$, where $k_j < k_l$ for $j < l$ (we drop index s for convenience); then $J \neq \varnothing$. In equilibrium, proposal P_{k_1} is accepted, so $\chi(s) = P_{k_1}$ and $\psi(P_{k_1}) = \psi(s)$. Since each P_{k_l} for $1 \leq l \leq |J|$ is accepted in this equilibrium, then $1 \leq l < |J|$, $\overline{V}_i(s, P_{k_l}) \geq \overline{V}_i(s, P_{k_{l+1}})$ for a winning coalition in s. For such players,

$$\text{(A10)} \qquad (1 - \beta^{m(P_{k_l})+1})\,\tilde{w}_i + \beta^{m(P_{k_l})+1} w_i(\psi(P_{k_l}))$$
$$\geq (1 - \beta^{m(P_{k_{l+1}})+1})\,\tilde{w}_i + \beta^{m(P_{k_{l+1}})+1} w_i(\psi(P_{k_{l+1}})),$$

and therefore, from (A3), $w_i(\psi(P_{k_l})) \geq w_i(\psi(P_{k_{l+1}}))$; this implies $\psi(P_{k_l}) \succeq_s \psi(P_{k_{l+1}})$. We also have that $\overline{V}_i(s, P_{k_{|J|}}) \geq \overline{V}_i(s, s)$ for a winning coalition in s, and for such players,

$$\text{(A11)} \quad \left(1 - \beta^{m(P_{k_{|J|}})+1}\right)\tilde{w}_i + \beta^{m(P_{k_{|J|}})+1} w_i(\psi(P_{k_{|J|}}))$$
$$\geq (1 - \beta)\, w_i(s) + \beta\left((1 - \beta^{m(s)})\,\tilde{w}_i + \beta^{m(s)} w_i(\psi(s))\right)$$
$$> (1 - \beta^{m(s)+1})\,\tilde{w}_i + \beta^{m(s)+1} w_i(\psi(s)).$$

From (A3), we get $w_i(\psi(P_{k_{|J|}})) \geq w_i(\psi(s)) = w_i(P_{k_1})$; therefore, $\psi(P_{k_{|J|}}) \succeq_s \psi(P_{k_1})$. Assumption 2(*ii*) now implies that $\psi(P_{k_j}) \sim \psi(P_{k_l})$ for all $1 \leq j < l \leq |J|$. Now (A10) implies that $m(P_{k_l}) \leq m(P_{k_{l+1}})$ for all $1 \leq l \leq |J| - 1$, which proves the auxiliary result.

Proof that ψ satisfies Axiom 1. Suppose $\psi(s) \neq s$, so the auxiliary result applies. For a winning coalition of players in s, $\overline{V}_i(s, P_{k_{|J|}}) \geq \overline{V}_i(s, s)$. The previous auxiliary result implies $\psi(P_{k_{|J|}}) = \psi(s)$ and $m(P_{k_1}) \leq m(P_{k_{|J|}}) = m(s) - 1$, and then the first inequality in (A11), together with (A3), implies $w_i(\psi(s)) > w_i(s)$. We have thus proved that for any $s \in \mathcal{S}$ such that $\psi(s) \neq s$, $\psi(s) \succ_s s$, and therefore Axiom 1 holds.

Proof that ψ satisfies Axiom 2 is straightforward as $\psi(\psi(s)) = \psi(s)$ from (A7).

Proof that $\chi(s) = \psi(s)$. If $\psi(s) = s$, then $\chi(s) = s = \psi(s)$ due to the "no cycle" result. Let us prove that if $\psi(s) \neq s$, then transition to state $\psi(s)$ takes place in one step; i.e., that $\psi(s) = \chi(s)$ (or, equivalently, in (A6) $m(s) = 1$ whenever $\chi(s) \neq s$). Consider two cases.

Case (i): $\psi(s) = P_{k_j}$ for some $j : 1 \leq j \leq |J|$. In this case, $m(P_{k_j}) = 0$ since Axiom 2 is proven to hold. But we proved that $m(P_{k_l})$ is weakly increasing in l, therefore, $m(\chi(s)) = m(P_{k_1}) = 0$, and therefore $m(s) = 1$.

Case (ii): $\psi(s) = P_{k_j}$ does not hold for any j. This implies that $m(P_{k_1}) \geq 1$ and $\psi(s) \neq \chi(s)$. Suppose that at some stage k, the proposal $P_k = \psi(s)$ is made (not necessarily on the equilibrium path). Then if it accepted, each player i will get $\overline{V}_i(s, P_k) = (1 - \beta)\,\tilde{w}_i + \beta w_i(\psi(s))$, and if it is rejected, he will receive

$$\overline{V}_i(s, x) \leq (1 - \beta)\, w_i(s) + \beta(1 - \beta)\,\tilde{w}_i + \beta^2 w_i(\psi(s))$$

for some x such that $\psi(x) = \psi(s)$. Any player with $w_i(\psi(s)) > w_i(s)$ must, given (A3), have $\overline{V}_i(s,P_k) > \overline{V}_i(s,x)$. Since $\psi(s) \succ_s s$ (Axiom 1), proposal $P_k = \psi(s)$ will be accepted.

By Assumption 5, either every proposal will be made exogenously at some stage k, or each player will become an agenda-setter. In the first case, $k \in J$, but in the case under consideration $\psi(s) = P_{k_j}$ does not hold for any j, yielding contradiction. In the second case, if a player i such that $w_i(\psi(s)) > w_i(s)$ is the agenda-setter at stage k, then he cannot propose $P_k = \psi(s)$ in equilibrium, as it will be accepted, and we again get to a contradiction. Proposing $P_k = \psi(s)$, however, will yield $\overline{V}_i(s,P_k)$ whereas making the equilibrium proposal will yield $\overline{V}_i(s,x)$. For player i, $\overline{V}_i(s,P_k) > \overline{V}_i(s,x)$ as we proved earlier, thus he has a profitable deviation. This cannot happen in equilibrium, which proves that $\chi(s) = \psi(s)$ for all $s \in \mathcal{S}$.

Proof that ψ satisfies Axiom 3. Suppose that Axiom 3 does not hold. This implies that there exist states $s, z \in \mathcal{S}$ such that $\psi(z) = z$, $z \succ_s s$ (which implies $z \neq s$), and $z \succ_s \psi(s)$ (which implies $\psi(z) \nsim \psi(s)$). As before, suppose that at some stage k, the proposal $P_k = z$ is made (not necessarily on equilibrium path). If it is accepted, each player i will get $\overline{V}_i(s,z) = (1 - \beta)\tilde{w}_i + \beta w_i(z)$, and if it is rejected, this player will get

$$\overline{V}_i(s,x) \leq (1 - \beta) w_i(s) + \beta(1 - \beta)\tilde{w}_i + \beta^2 w_i(\psi(s))$$

for some x such that $\psi(x) = \psi(s)$. Now, (A4) implies that $\overline{V}_i(s,z) > \overline{V}_i(s,x)$ whenever $w_i(z) > w_i(\psi(s))$; i.e., for a winning coalition in s. Therefore, proposal $P_k = z$ will be accepted.

Since $\psi(z) \nsim \psi(s)$, it must be that z is never proposed along the equilibrium path. By Assumption 5, this is only possible if each player becomes the agenda-setter at some stage k. When a player with $w_i(z) > w_i(\psi(s))$ becomes the agenda-setter, proposing z is a profitable deviation for him. This cannot happen in equilibrium, and this contradiction establishes that ψ satisfies Axiom 3. This completes the proof of part 2 of the Theorem.

Part 3: This result immediately follows from Theorem 1 and part 2 of this Theorem.

PROOF OF THEOREM 3:

Suppose, to obtain a contradiction, that stable state $s \in \mathcal{S}$ is Pareto inefficient. This means that for some $x \in \mathcal{S}$, $w_i(x) > w_i(s)$ for all $i \in \mathcal{I}$. By hypothesis, there is $y \in \mathcal{S}$ such that $\mathcal{W}_y \subset \mathcal{W}_s$ and $w_i(y) = w_i(x) > w_i(s)$ for all $i \in \mathcal{I}$. Take a mapping $\phi \in \Phi$. that satisfies Axioms 1–3. Consider two cases. If $\phi(y) = y$, then from $\phi(s) = s$ and $y \succ_s s$ we get $y \succ_s \phi(s)$, ϕ violates Axiom 3 (if there is z such that $\phi(y) = y$, $y \succ_s s$, and $y \succ_s z$, then $z \neq \phi(s)$). If $\phi(y) \neq y$, then Axiom 1 implies $w_i(\phi(y)) > w_i(y) > w_i(s)$ for a winning coalition in y, which is a winning coalition in s, and thus $\phi(y) \succ_s s$ and $\phi(y) \succ_s \phi(s)$. Axiom 2 guarantees that $\phi(\phi(y)) = \phi(y)$. Again, we conclude that ϕ violates Axiom 3.

PROOF OF THEOREM 4:

The next lemma, proved in the online Appendix, characterizes properties of quasi-median voters. Recall that M_s denotes the set of quasi-median voters in state s.

LEMMA 1: *Given* $\mathcal{I} \subset \mathbb{R}$, $\mathcal{S} \subset \mathbb{R}$, *payoff functions* $\{w_i(s)\}_{i \in \mathcal{I}, s \in \mathcal{S}}$, *and winning coalitions* $\{\mathcal{W}_s\}_{s \in \mathcal{S}}$ *satisfying Assumption* 1, *the following are true.*

(i) *For each s, the set* M_s *is nonempty.*

(ii) *If the single-crossing property in Definition 3 holds, then for any states* $x, y, z \in \mathcal{S}$,

$$x \succ_z y \text{ if and only if for all } i \in M_z, \quad w_i(x) > w_i(y), \text{ and}$$

$$x \succeq_z y \text{ if and only if for all } i \in M_z, \quad w_i(x) > w_i(y).$$

(iii) *If monotonic median voter condition in Definition 5 holds, then there is a nondecreasing sequence* $\{m_s\}_{s \in \mathcal{S}}$ *of players such that* $m_s \in M_s$ *for all* $s \in \mathcal{S}$.

PROOF OF THEOREM 4:

Part 1: We start with Assumption 2(i). Suppose that there is a cycle $s_1, \ldots, s_l$ such that $s_{k+1} \succ_{s_k} s_k$ for $1 \leq k \leq l-1$ and $s_1 \succ_{s_l} s_l$. Take a monotonic sequence of median voters $\{m_s\}_{s \in \mathcal{S}}$. Recall that m_s is part of any connected winning coalition in s, therefore, if for some x and z, $x \succ_z z$, then $w_x(m_z) > w_z(m_z)$. Now for each $s \in \mathcal{S}$ consider an alternative set of winning coalitions where m_s is the dictator; i.e., $\mathcal{W}'_s = \{X \in \mathcal{C} : m_s \in X\}$. Denoting the induced relation between states by $\succ'$, we have that if $x \succ_z z$, then $x \succ'_z z$. Consequently, if there was a cycle $s_1, \ldots, s_l$ such that $s_{k+1} \succ_{s_k} s_k$ for $1 \leq k \leq l-1$ and $s_1 \succ_{s_l} s_l$, then we have $s_{k+1} \succ'_{s_k} s_k$ for $1 \leq k \leq l-1$ and $s_1 \succ'_{s_l} s_l$; therefore, a cycle for $\succ'$ exists. Now take the shortest cycle for $\succ'$ (which need not be a cycle for $\succ$). Without loss of generality, suppose that s_2 is the lowest state (so $s_2 \leq s_1$ and $s_2 \leq s_3$); then $m_{s_2} \leq m_{s_1}$ and $m_{s_2} \leq m_{s_3}$. Since $s_3 \succ'_{s_2} s_2$ and $s_2 \succ'_{s_1} s_1$, we have $w_{m_{s_2}}(s_3) > w_{m_{s_2}}(s_2)$ and $w_{m_{s_1}}(s_2) > w_{m_{s_1}}(s_1)$. But $s_2 \leq s_3$ and $m_{s_2} \leq m_{s_1}$, hence, $w_{m_{s_2}}(s_3) - w_{m_{s_2}}(s_2) > 0$ implies $w_{m_{s_1}}(s_3) - w_{m_{s_1}}(s_2) > 0$. Combining this with $w_{m_{s_1}}(s_2) > w_{m_{s_1}}(s_1)$, we conclude that $w_{m_{s_1}}(s_3) > w_{m_{s_1}}(s_1)$. But then $s_3 \succ'_{s_1} s_1$, since m_{s_1} is the dictator in s_1. This implies that s_2 may be skipped in the cycle, contradicting the assumption that $\{s_k\}_{k=1}^{l}$ is the shortest cycle.

To verify Assumption 2(ii), take any $s \in \mathcal{S}$ and some $m_s \in M_s$. Suppose there is a cycle $s_1, \ldots, s_l$ such that $s_{k+1} \succeq_s s_k$ for $1 \leq k \leq l-1$, $s_1 \succeq_s s_l$, and $s_j \nsim s_k$ for $1 \leq j < k \leq l$. Without loss of generality, assume that state s_l maximizes the payoff of m_s among states $s_1, \ldots, s_l$. Then $w_{m_s}(s_l) \geq w_{m_s}(s_1)$, and Assumption 6 implies $w_{m_s}(s_l) > w_{m_s}(s_1)$. But then, by Lemma 1, $s_1 \not\succeq_s s_l$, and this contradicts the existence of a cycle. Finally, if $x, y \in \mathcal{S}$ are such that $x \succ_s s$ and $y \succ_s x$, then for any $i \in M_s$ we have $w_i(y) > w_i(x) > w_i(s)$, which, in turn, implies $y \succ_s s$. This shows that Assumption 2(ii) holds and completes the proof of part 1.[24]

[24] This result can also be derived using Theorem 4.6 in Austin-Smith and Banks (1999).

Part 2:[25] Let $\mathcal{W} = \cup_{s \in \mathcal{S}} \mathcal{W}_s$; then $\mathcal{W}$, as a set of winning coalitions, satisfies Assumption 1. Let $\succ^*$ be given by $x \succ^* y$ if and only if $\{i \in \mathcal{I} : w_i(x) > w_i(y)\} \in \mathcal{W}$. Since preferences are single-peaked, Theorem 4.1 in Austen-Smith and Banks (1999) implies that $\succ^*$ is transitive, and hence acyclic. Clearly, a cycle in Assumption 2(i) would also be a cycle for $\succ^*$; given Assumption 6, so would a cycle in Assumption 2(ii). Hence, such cycles do not exist. Finally, Theorem 4.1 in Austen-Smith and Banks (1999) suggests that the preference relation $\succ_s$ is transitive, and so $x \succ_s s$ and $y \succ_s x$ imply $y \succ_s s$.

Proofs of Propositions in Section V

PROOF OF PROPOSITION 1:

Part 1: Since $\alpha > \frac{1}{2}$ is the rule for all states, Assumption 1 is satisfied. Enumerate all states as $s_1, \ldots, s_m$ (where $m = |\mathcal{S}|$) such that ρ_{s_k} is weakly increasing in k (the order of states with the same ρ may be arbitrary). With this order, $\mathcal{I}$ and $\mathcal{S}$ satisfy the single-crossing condition as in Definition 3. Indeed, if $s_k < s_l$ and $i < j$, then

$$\left(w_j(s_l) - w_j(s_k)\right) - \left(w_i(s_l) - w_i(s_k)\right)$$

$$= \left(v_j(\rho_{s_l}) - v_j(\rho_{s_k})\right) - \left(v_i(\rho_{s_l}) - v_i(\rho_{s_k})\right) \geq 0,$$

because $\rho_{s_k} \leq \rho_{s_l}$ and v satisfies the strict increasing differences condition. Now construct a nondecreasing sequence of quasi-median voters; this would prove that monotonic median voter property holds. For state s_k, take m_{s_k} such that $\hat{\rho}_{m_{s_k}} \leq \rho_{s_k} < \hat{\rho}_{m_{s_k}+1}$ if $\rho_{s_k} < \hat{\rho}_n$, and let $m_{s_k} = n$ otherwise. Then m_{s_k} is determined uniquely for each state s_k, is weakly increasing, and is a quasi-median voter in state s_k by the assumption on feasible religious policies ρ. We can now apply part 1 of Theorem 4 to show that Assumption 2 is satisfied and Theorem 1 and Theorem 2 apply.

Part 2: Suppose that some state s with $1, n \notin Z(s)$ is stable. Suppose $Z(s) = [a, b]$ and let $Z' = [a - 1, b + 1]$. Then $\min M_{Z'} \leq \min M_Z \leq \max M_Z \leq \max M_{Z'}$, and thus $s' = (\rho_s, Z')$ is a feasible state. By the assumption on $V(Z)$, s' Pareto dominates s. Take a mapping ϕ that satisfies Axioms 1–3 and let $x = \phi(s')$. Then Axiom 1 implies that $w_i(\phi(s')) \geq w_i(s') > w_i(s)$ for a winning coalition in s', and thus, by Lemma 1, for all $i \in M_{s'}$. Therefore, this holds for all $i \in M_s$, and thus for a winning coalition in s. Since $\phi(\phi(s')) = \phi(s')$, we get a violation of Axiom 3. This proves that s is not stable, and either 1 or n should be part of $Z(s)$ for any stable state s. Hence, starting from any s_0, at least one of these players will be given political participation.

[25] We thank an anonymous referee for suggesting this simpler proof of part 2 of Theorem 4.

PROOF OF PROPOSITION 2:

Part 1: Notice that any state x with $Z(x) = \mathcal{I}$ is stable, as any ϕ with $\phi(x) \neq x$ would violate Axiom 1. Indeed, since $\hat{\rho}_{\min M_{\mathcal{I}}} \leq \rho_x \leq \hat{\rho}_{\max M_{\mathcal{I}}}$ and preferences are single-peaked, quasi-median voter $\min M_{\mathcal{I}}$ would be worse off from any state y with $\rho_y > \rho_x$, and max $M_{\mathcal{I}}$ would be worse off if $\rho_y < \rho_x$. Now suppose, to obtain a contradiction, that for some s such that $Z(s) \neq \mathcal{I}$, $\phi(s) = s$. Consider the following cases. Case (i): $\rho(s) < \hat{\rho}_{\min M_{\mathcal{I}}}$. Take $x = (\hat{\rho}_{\min M_{\mathcal{I}}}, \mathcal{I})$; by hypothesis, $w_1(x) > w_1(s)$, and thus

$$w_1(x) = v_1(\hat{\rho}_{\min M_{\mathcal{I}}}) + V(\mathcal{I}) > v_1(\hat{\rho}_1) + V(s) \geq v_1(s) + V(s) = w_i(s).$$

Since $\rho(x) > \rho(s)$, this implies $w_i(x) > w_i(s)$ for all $i \in \mathcal{I}$. But we proved that x is stable, and then $\phi(s) = s$ violates Axiom 3. Case (ii): $\hat{\rho}_{\min M_{\mathcal{I}}} \leq \rho_s \leq \hat{\rho}_{\max M_{\mathcal{I}}}$. Take $x = (\rho_s, \mathcal{I})$ and notice that $w_i(x) > w_i(s)$ for all $i \in \mathcal{I}$. Since we earlier proved that $\phi(x) = x$, we immediately get a contradiction to Axiom 3. Case (iii): $\rho_s > \hat{\rho}_{\max M_{\mathcal{I}}}$. This case is completely analogous to case (i). In all cases, $\phi(s) = s$ leads to a contradiction.

Part 2: Let $k = 1$. Let state s be such that $Z(s) \leq k$ and $Z(s)$ includes the middle player. Denote the player in $Z(s)$ by i and take $x = (\hat{\rho}_i, \mathcal{I})$, which is feasible for any α. Then $\phi(x) = x$ as proved earlier. If $\phi(s) \neq x$, then Axiom 3 is violated, as $w_i(x) > w_i(y)$ for any state $y \neq x$. This proves that $\phi(s) = x$.

If s includes either player 1 or player n and $Z(s) \leq k$, then either $s = (\hat{\rho}_1, \{1\})$ or $s = (\hat{\rho}_n, \{n\})$. Suppose that $\phi(s) = x = (\rho, \mathcal{I})$; then $\hat{\rho}_{\min M_{\mathcal{I}}} \leq \rho \leq \hat{\rho}_{\max M_{\mathcal{I}}}$, and $v_1(\rho) \leq v_1(\hat{\rho}_{\min M_{\mathcal{I}}})$ and $v_n(\rho) \leq v_1(\hat{\rho}_{\max M_{\mathcal{I}}})$ by single-peakedness. Then in the first case,

$$w_1(x) = v_1(\rho) + V(\mathcal{I}) \leq v_1(\hat{\rho}_{\min M_{\mathcal{I}}}) + V(\mathcal{I}) < v_1(\hat{\rho}_1) + V(\{1\}) = w_1(s),$$

and thus Axiom 1 is violated. Similarly, in the latter case, $w_n(x) < w_1(s)$, and Axiom 1 is again violated. This proves that $\phi(s) \neq x$.

Part 3: Take some s; suppose, to obtain a contradiction, that $\phi(s) = x$ such that $Z(x) \neq \mathcal{I}$. By Axiom 1, $w_i(x) > w_i(s)$ for all $i \in \mathcal{I}$. Consider $y = (\rho_x, \mathcal{I})$; the unanimity rule ensures that y is feasible for any $\rho(x)$. As shown earlier, $\phi(y) = y$, and $w_i(y) > w_i(x) > w_i(s)$ for every player $i \in \mathcal{I}$. But then $\phi(s) = x$ violates Axiom 3, a contradiction.

PROOF OF PROPOSITION 3:

Part 1: The proof follows that of part 1 of Proposition 1.

Part 2: Suppose, to obtain a contradiction, that for some s, $\phi(s) = x$ such that $Z(x) \neq \mathcal{I}$ for a mapping ϕ that satisfies Axioms 1 – 3. By Axiom 1, $w_i(x) > w_i(s)$ for a winning coalition in s. Consider $y = (\alpha', \rho_x, \mathcal{I})$ such that $\alpha' > \frac{n-1}{n}$; then y is feasible. But $\phi(y) = y$, and $w_i(y) > w_i(x)$ for every player $i \in \mathcal{I}$. Then $w_i(y) > w_i(s)$ for a winning coalition in x. But then $\phi(s) = x$ violates Axiom 3, a contradiction.

REFERENCES

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2008. "Coalition Formation in Non-Democracies." *Review of Economic Studies* 75 (4): 987–1009.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2009. "Equilibrium Refinement in Dynamic Voting Games." Massachusetts Institute of Technology Department of Economics Working Paper 09-26.

Acemoglu, Daron, and James Robinson. 2000. "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective." *Quarterly Journal of Economics* 115 (4): 1167–99.

Acemoglu, Daron, and James Robinson. 2006. *Economic Origins of Dictatorship and Democracy*. New York: Cambridge University Press.

Alesina, Alberto, Ignazio Angeloni, and Federico Etro. 2005. "International Unions." *American Economic Review* 95 (3): 602–15.

Austen-Smith, David, and Jeffrey S. Banks. 1999. *Positive Political Theory I: Collective Preference*. Ann Arbor, MI: University of Michigan Press.

Barberà, Salvador, and Matthew O. Jackson. 2004. "Choosing How to Choose: Self-Stable Majority Rules and Constitutions." *Quarterly Journal of Economics* 199 (3): 1011–48.

Barberà, Salvador, Michael Maschler, and Jonathan Shalev. 2001. "Voting for Voters: A Model of the Electoral Evolution." *Games and Economic Behavior* 37 (1): 40–78.

Barberà, Salvador, Hugo Sonnenschein, and Lin Zhou. 1991. "Voting by Committees." *Econometrica* 59 (3): 595–609.

Bordignon, Massimo, and Sandro Brusco. 2006. "On Enhanced Cooperation." *Journal of Public Economics* 90 (10–11): 2063–90.

Bourguignon, François, and Thierry Verdier. 2000. "Oligarchy, Democracy, Inequality and Growth." *Journal of Development Economics* 62 (2): 285–313.

Burkart, Mike, and Klaus Wallner. 2000. "Club Enlargement: Early versus Late Admittance." http://ssrn.com/abstract=248994.

Chwe, Michael S. Y. 1994. "Farsighted Coalitional Stability." *Journal of Economic Theory* 63: 299–325.

Ellickson, Bryan, Birgit Grodal, Suzanne Scotchmer, and William R. Zame. 1999. "Clubs and the Market." *Econometrica* 67 (5): 1185–1217.

Gans, Joshua, and Michael Smart. 1996. "Majority Voting with Single-Crossing Preferences." *Journal of Public Economics* 59 (2): 219–37.

Gomes, Armando, and Philippe Jehiel. 2005. "Dynamic Processes of Social and Economic Interactions: On the Persistence of Inefficiencies." *Journal of Political Economy* 113 (3): 626–67.

Jehiel, Philippe, and Suzanne Scotchmer. 2001. "Constitutional Rules of Exclusion in Jurisdiction Formation." *Review of Economic Studies* 68: 393–413.

Konishi, Hideo, and Debraj Ray. 2003. "Coalition Formation as a Dynamic Process." *Journal of Economic Theory* 110 (1): 1–41.

Lagunoff, Roger. 2009. "The Dynamic Stability and Reform of Political Institutions." *Games and Economic Behavior* 67: 569–83.

Lizzeri, Alessandro, and Nicola Persico. 2004. "Why Did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's 'Age of Reform.'" *Quarterly Journal of Economics* 119 (2): 705–63.

Mariotti, Marco. 1997. "A Model of Agreements in Strategic Form Games." *Journal of Economic Theory* 74 (1): 196–217.

Rabasa, Angel, and Stephen Larrabee. 2008. *The Rise of Political Islam in Turkey.* Santa Monica, CA: RAND.

Ray, Debraj, and Rajiv Vohra. 1997. "Equilibrium Binding Agreements." *Journal of Economic Theory* 73 (1): 30–78.

Ray, Debraj, and Rajiv Vohra. 1999. "A Theory of Endogenous Coalition Structures." *Games and Economic Behavior* 26 (2): 286–336.

Roberts, Kevin W. S. 1999. "Dynamic Voting in Clubs." http://ssrn.com/abstract=1160966.

Rothstein, Paul. 1990. "Order Restricted Preferences and Majority Rule." *Social Choice and Welfare* 7 (4): 331–42.

Roy, Olivier. 2007. *Secularism Confronts Islam.* New York: Columbia University Press.

Schauer, Frederick. 1985. "Slippery Slopes." *Harvard Law Review* 99 (2): 361–83.

Scotchmer, Suzanne. 2002. "Local Public Goods and Clubs." In *Handbook of Public Economics*, edited by Alan Auerbach and Martin Feldstein, 1997–2042. Amsterdam: North-Holland.

Seidmann, Daniel, and Eyal Winter. 1998. "A Theory of Gradual Coalition Formation." *Review of Economic Studies* 65 (4): 793–815.