Online Appendix to
"Who Controls the Agenda Controls the Legislature"

S. Nageeb Ali **B. Douglas Bernheim** **Alexander W. Bloedel**
Silvia Console Battilana

This Online Appendix presents proofs and examples omitted from the main paper. We present this material in the same order it is mentioned in the main text. All references to the main paper follow the numbering conventions therein. The numbering of equations, figures, and results introduced in this Online Appendix begins where the corresponding numbering in the main paper ends.

B Omitted Proofs and Examples

B.1 Proof of Lemma 3

We consider each direction in turn. For the "if" direction, suppose $\mathcal{C}$ satisfies Thin Individual Indifference. Let $x \in X$ and $\epsilon > 0$ be given. Note that because X is compact, the open covering $\left\{B_{\epsilon/2}(y)\right\}_{y\in X}$ has a finite subcovering; enumerate it by $\{B_k\}_{k=1}^K$ and suppose, without loss of generality, that $x \in B_1$. For any $y \in X$, let $\mathcal{I}(y) := \bigcup_{i\in N\cup\{A\}} I_i(y)$. Recursively construct the sequences $\{D_k\}_{k=1}^{K-1} \subseteq 2^X$ and $\{x_k\}_{k=1}^K \subseteq X$ as follows:

- Let $D_0 := \emptyset$ and $D_k := D_{k-1} \bigcup [\mathcal{I}(x_k)\backslash\{x_k\}]$ for $k \geq 1$;
- Let $x_1 := x \in B_1$ and pick $x_k \in B_k\backslash D_{k-1}$ arbitrarily for $k \geq 2$.

We claim that $X_\epsilon := \{x_k\}_{k=1}^K$ is a generic ϵ-grid; since $x \in X_\epsilon$ by construction, this suffices to prove that $\mathcal{C}$ is Finitely Approximable. We establish the claim in three steps.

Step 1: The sequences $\{D_k\}_{k=1}^{K-1}$ and $\{x_k\}_{k=1}^K$ are well-defined, viz., $B_k\backslash D_{k-1} \neq \emptyset$ for all k. The argument is by induction. For the base step, note that D_1 has empty interior by Thin Individual Indifference. Because B_2 is nonempty and open, it then follows that $B_2\backslash D_1 \neq \emptyset$ and therefore x_2 is well-defined. For the inductive step, let $2 \leq k \leq K-1$ be given and suppose that D_{j-1} and x_j are well-defined for all $j \leq k$.

We assert that D_k has empty interior. To prove this, for each $j \leq k$ we define

$$\Delta_j := \begin{cases} \mathcal{I}(x_j) \backslash \{x_j\} & \text{if } x_j \text{ is an isolated point,} \\ \mathcal{I}(x_j) & \text{otherwise.} \end{cases} \tag{9}$$

We show that each Δ_j is closed and has empty interior, considering the two cases in (9) in turn. First, suppose that x_j is isolated, i.e., $\{x_j\}$ is open. Since $\mathcal{I}(x_j)$ is closed (by continuity of preferences), this implies that $\Delta_j = \mathcal{I}(x_j) \backslash \{x_j\}$ is closed. Meanwhile, Thin Individual Indifference implies that Δ_j has empty interior. Second, suppose that x_j is not isolated, i.e., $\{x\}$ is not open. Then $\Delta_j = \mathcal{I}(x_j)$ is closed (again by continuity of preferences). Towards a contradiction, suppose that Δ_j has nonempty interior, i.e., there exists some $y \in \Delta_j$ and $\delta > 0$ such that $B_\delta(y) \subseteq \Delta_j$. If $y \neq x_j$, then $d(y, x_j) > 0$; picking $\delta \in (0, d(y, x_j))$ implies that $B_\delta(y) \subseteq \mathcal{I}(x_j) \backslash \{x_j\}$, contradicting Thin Individual Indifference. If $y = x_j$, then $O := B_\delta(x_j) \backslash \{x_j\}$ is a nonempty open set (because x_j is not isolated) and $O \subseteq \mathcal{I}(x_j) \backslash \{x_j\}$ by construction, again contradicting Thin Individual Indifference. We conclude that each Δ_j is closed and has empty interior, as desired. It then follows that $E_k := \bigcup_{j=1}^k \Delta_j$ also has empty interior. Since, by construction, $D_k = \bigcup_{j=1}^k [\mathcal{I}(x_j) \backslash \{x_j\}]$, and $\mathcal{I}(x_j) \backslash \{x_j\} \subseteq \Delta_j$ for all $j \leq k$, we obtain that $D_k \subseteq E_k$. Therefore, D_k has empty interior, as desired.

As in the base step, it then follows that $B_{k+1} \backslash D_k \neq \emptyset$ and therefore x_{k+1} is well-defined. This completes the induction.

Step 2: All players have strict preferences on X_ϵ. As $X_\epsilon := \{x_k\}_{k=1}^K$, it suffices to show:[1]

$$\text{For all } \ell, k \in \{1, \ldots, K\} \text{ with } \ell \leq k,\ x_k \in \mathcal{I}(x_\ell) \text{ implies that } x_k = x_\ell. \tag{10}$$

To this end, note that, by construction, for all $\ell \leq k$ we have

$$x_k \notin D_{k-1} = \bigcup_{j=1}^{k-1} [\mathcal{I}(x_j) \backslash \{x_j\}] \supseteq \bigcup_{j=1}^{\ell-1} [\mathcal{I}(x_j) \backslash \{x_j\}] \supseteq \left[\bigcup_{j=1}^{\ell-1} \mathcal{I}(x_j) \right] \backslash \{x_1, \ldots, x_{\ell-1}\}.$$

In turn, this implies the nontrivial ("only if") direction of the following property:

$$\text{For all } \ell \leq k,\ x_k \in \bigcup_{j=1}^{\ell-1} \mathcal{I}(x_j) \text{ if and only if } x_k \in \{x_j\}_{j=1}^{\ell-1}. \tag{11}$$

For each k, let $\rho(k) := \min \left\{ \hat{k} \in \mathbb{N} : x_k \in \bigcup_{j=1}^{\hat{k}} \mathcal{I}(x_j) \right\}$. Note that $\rho(k) \leq k$ and, by (11), $x_k = x_{\rho(k)}$.

[1] In (10), it is without loss of generality to let $\ell \leq k$ because $x_k \in \mathcal{I}(x_\ell)$ if and only if $x_\ell \in \mathcal{I}(x_k)$.

We now prove (10) by induction. For the base ($k = 2$) step, (11) implies that $x_2 \in \mathcal{I}(x_1)$ only if $x_2 = x_1$. For the inductive step, let $k \in \{2, \ldots, K-1\}$ be given and suppose that, for all $m \leq \ell \leq k$, $x_\ell \in \mathcal{I}(x_m)$ only if $x_\ell = x_m$. We claim that, for all $\hat{k} \leq k+1$, $x_{k+1} \in \mathcal{I}(x_{\hat{k}})$ only if $x_{k+1} = x_{\hat{k}}$. This holds for all $\hat{k} \leq \rho(k+1)$ and $\hat{k} = k+1$ by construction, so suppose that $x_{k+1} \in \mathcal{I}(x_{\hat{k}})$ for some $\hat{k} \in \{\rho(k+1)+1, \ldots, k\}$. As $x_{k+1} = x_{\rho(k+1)}$, we have $x_{\rho(k+1)} \in \mathcal{I}(x_{\hat{k}})$, which implies that $x_{\hat{k}} \in \mathcal{I}(x_{\rho(k+1)})$. As $\rho(k+1) \leq \hat{k} \leq k$, the inductive hypothesis implies that $x_{\hat{k}} = x_{\rho(k+1)} = x_{k+1}$. This completes the induction.

Step 3: X_ϵ is an ϵ-grid, viz., $\max_{y \in X} d(y, X_\epsilon) < \epsilon$. Recall that $\{B_k\}_{k=1}^K$ is a covering of X by open balls of radius $\epsilon/2$, while X_ϵ is constructed from a selection $k \mapsto x_k \in B_k$. Hence, $\sup_{y \in B_k} d(y, x_k) < \epsilon$ for every $k \leq K$. Finally, observe that

$$\max_{y \in X} d(y, X_\epsilon) \leq \max_{k \leq K} \left[\sup_{y \in B_k} d(y, x_k) \right] < \epsilon.$$

This completes our proof of the sufficiency of Thin Individual Indifference.

To establish its necessity—the "only if" direction of the lemma—suppose that $\mathcal{C}$ violates Thin Individual Indifference. Then there exists a policy $x \in X$, player $i \in N \cup \{A\}$, and nonempty open set $O \subset X$ such that $O \subset I_i(x) \backslash \{x\}$. Pick $y \in O$ and $\delta > 0$ so that $B_\delta(y) := \{z \in X : d(y, z) < \delta\} \subseteq O$. Given any $\epsilon \in (0, \delta]$, let $X_\epsilon \subset X$ be a (not necessarily generic) ϵ-grid for which $x \in X_\epsilon$. There exists some $z \in X_\epsilon \cap B_\delta(y)$ by definition of X_ϵ,[2] and hence $z \in X_\epsilon \cap I_i(x) \backslash \{x\}$ by definition of $B_\delta(y)$, implying that player i's preferences are not strict on X_ϵ. It follows that $\mathcal{C}$ does not admit any generic ϵ-grid containing x with $\epsilon \in (0, \delta]$, and therefore is not Finitely Approximable. □

B.2 Details and Proofs for Theorem 3

We provide the main proof of Theorem 3 in Appendix B.2.1. The key lemmas presented therein, Lemmas 5 and 6, are proved separately in Appendices B.2.2 and B.2.3.

B.2.1 Proof of Theorem 3

Step 1: Preliminaries. We define the *weak* majority acceptance correspondence by $M^{\mathrm{w}}(x) := \{y \in X : y \succcurlyeq_M x\}$, the *strict* majority acceptance correspondence by $M^{\mathrm{s}}(x) := \{y \in X : y \succ_M x\}$, and the *almost-strict* majority acceptance correspondence by $M^{\mathrm{as}}(x) := \mathrm{cl}\,[M^{\mathrm{s}}(x)] \cup \{x\}$. Define the agenda setter's *favorite almost-strict improvement* value function $V_A^{\mathrm{as}} : X \to \mathbb{R}$ by $V_A^{\mathrm{as}}(x) := \max_{y \in M^{\mathrm{as}}(x)} u_A(y)$, and her *one-round improvement* correspondence $\Phi^{\mathrm{or}} : X \rightrightarrows X$

[2] If not, then $\min_{z \in X_\epsilon} d(y, z) \geq \delta$ and so X_ϵ would not be an ϵ-grid with $\epsilon \leq \delta$.

by

$$\Phi^{\text{or}}(x) := \{y \in X : y \in M^{\text{w}}(x) \text{ and } u_A(y) \geq V_A^{\text{as}}(x)\}. \tag{12}$$

Because the policy space is compact and all players' preferences are continuous, each correspondence described above is nonempty- and compact-valued. We denote the set of unimprovable policies—as in Definition 1—by $\mathcal{E}$.

In settings with Generic Finite Alternatives, $\Phi^{\text{or}}(x) = \{\phi(x)\}$, where ϕ is the favorite improvement function defined in Equation (1). Lemma 4 shows that Φ^{or} is the appropriate generalization of ϕ to general collective choice problems.

Lemma 4. *For any collective choice problem $\mathcal{C}$, the following hold:*

(a) *The set of unimprovable policies satisfies $\mathcal{E} = \{x \in X : x \in \Phi^{or}(x)\}$.*

(b) *For any pair of policies x and y, we have $y \in \Phi^{or}(x)$ if and only if y is the outcome of some Non-Capricious equilibrium of the one-round game with initial default x.*

Proof. To prove part (a), let $x \in X$ be given. We show that $x \in \mathcal{E}$ if and only if $x \in \Phi^{\text{or}}(x)$. For the "if" direction, note that $x \in \Phi^{\text{or}}(x)$ implies that $u_A(x) \geq V_A^{\text{as}}(x)$, and hence there do not exist any $y \in M^{\text{as}}(x)$ such that $y \succ_A x$; as $M^{\text{s}}(x) \subseteq M^{\text{as}}(x)$, it follows that $x \in \mathcal{E}$. For the "only if" direction, suppose that $x \notin \Phi^{\text{or}}(x)$. Then, because $x \in M^{\text{w}}(x)$, it must be that $u_A(x) < V_A^{\text{as}}(x)$. Thus, there exists some $y \in M^{\text{as}}(x) \backslash \{x\}$ such that $y \succ_A x$. Because $M^{\text{as}}(x) \backslash \{x\} = \text{cl}\,[M^{\text{s}}(x)] \backslash \{x\}$ by definition, there exists a sequence $\{y^n\} \subseteq M^{\text{s}}(x)$ such that $y^n \to y$ and, being that $\succcurlyeq_A$ is continuous, there exists an $N \in \mathbb{N}$ such that $y^n \succ_A x$ for all $n \geq N$. Thus, any y^n with $n \geq N$ is an improvement to x, implying that $x \notin \mathcal{E}$.

To prove part (b), let $x, y \in X$ be given. For the "if" direction, suppose that y is the outcome under a Non-Capricious equilibrium σ in the one-round game with default x. Suppose, towards a contradiction, that $y \notin \Phi^{\text{or}}(x)$. We first establish that $y \neq x$. Suppose otherwise. Then by part (a), x is improvable, which implies that there exists some $z \in M^{\text{s}}(x)$ such that $z \succ_A x$. This is incompatible with the hypothesis that σ is an equilibrium: because voters would pass this z with probability one, the agenda setter could profitably deviate by proposing z. Having established that $y \neq x$, it must be that y is proposed and accepted with probability one under σ. There are two cases. First, if $y \notin M^{\text{w}}(x)$, then there exists some voter $i \in N$ such that i votes to approve proposal y under σ, and yet $x \succ_i y$; voter i then has a strictly profitable deviation from voting to reject y. Second, suppose that $u_A(y) < V_A^{\text{as}}(x)$. By definition, $V_A^{\text{as}}(x) \geq u_A(x)$. If $V_A^{\text{as}}(x) = u_A(x)$, then the agenda setter has a strictly profitable deviation from proposing x instead of y, as this implements x regardless of the voters' response and $u_A(x) > u_A(y)$. If $V_A^{\text{as}}(x) > u_A(x)$, then $V_A^{\text{as}}(x) = \max_{z \in \text{cl}[M^s(x)] \backslash \{x\}} u_A(z) = \sup_{z \in M^s(x)} u_A(z)$, where the second equality follows from the continuity of u_A. Thus, there exists some $z \in M^{\text{s}}(x)$ such that $u_A(z) > u_A(y)$. The agenda setter then has a strictly profitable deviation from proposing z instead of y, as every policy in $M^{\mathrm{s}}(x)$ must be accepted by a majority of voters in every equilibrium. In either case, we establish that σ is not an equilibrium, obtaining the desired contradiction.

For the "only if" direction of part (b), suppose that $y \in \Phi^{\mathrm{or}}(x)$. Consider the following pure strategy profile in the one-round game with initial default x: the agenda setter proposes y and each voter $i \in N$ votes to accept a proposal z if and only if either (i) $z \succ_i x$ or (ii) $z \sim_i x$ and $z = y$. By construction, no voter has a profitable deviation; the agenda setter's payoff from proposing y is $u_A(y) \geq V_A^{\mathrm{as}}(x)$, while her payoff from any other proposal is bounded above by $\max\{u_A(x), \sup_{z \in M^{\mathrm{s}}(x)} u_A(z)\} \leq V_A^{\mathrm{as}}(x)$, so that she also has no profitable deviations. Therefore, this strategy profile is an equilibrium. As every pure-strategy equilibrium of any one-round game is Non-Capricious, this strategy profile is a Non-Capricious equilibrium inducing outcome y. □

Step 2: Non-Capricious Equilibrium Outcomes. We now characterize non-capricious equilibrium outcomes in general collective choice problems. Let $\Sigma^{\mathrm{NC}}(x^0, T)$ denote the set of non-capricious equilibria of the game with T rounds and initial default x^0. Let $g_T^\sigma(x^0) \in X$ denote the outcome induced by equilibrium $\sigma \in \Sigma^{\mathrm{NC}}(x^0, T)$, and $G_T(x^0) := \bigcup_{\sigma \in \Sigma^{\mathrm{NC}}(x^0,T)} \{g_T^\sigma(x^0)\}$ denote those across all non-capricious equilibria.

We characterize outcomes for all equilibria using the Φ^{or} operator. We say that $\hat{\phi} : X \to X$ is a *selection* of Φ^{or} if $\hat{\phi}(x) \in \Phi^{\mathrm{or}}(x)$ for every $x \in X$; we denote selections by $\hat{\phi}(\cdot) \in \Phi^{\mathrm{or}}(\cdot)$.

Lemma 5. *For any collective choice problem $\mathcal{C}$, $x^0 \in X$, and $T \in \mathbb{N}$, the following hold:*

(a) *For any selection $\hat{\phi}(\cdot) \in \Phi^{or}(\cdot)$, there exists a Non-Capricious equilibrium $\sigma \in \Sigma(x^0, T)$ inducing the outcome $g_T^\sigma(x^0) = \hat{\phi}^T(x^0)$.*

(b) *For any Non-Capricious equilibrium $\sigma \in \Sigma^{NC}(x^0, T)$, there exists a collection $\{\hat{\phi}_t(\cdot)\}_{\tau=1}^T$ of selections $\hat{\phi}_t(\cdot) \in \Phi^{or}(\cdot)$ such that the equilibrium outcome is given by*

$$g_T^\sigma(x^0) = \left[\hat{\phi}_1 \circ \hat{\phi}_2 \circ \cdots \circ \hat{\phi}_T\right](x^0)$$

and, for every $x \in X$, we have $\hat{\phi}_t(x) \sim_A \hat{\phi}_T(x)$ for all $1 \leq t \leq T$.

To interpret Lemma 5, observe that in the special case where Generic Finite Alternatives holds, it reduces to the characterization from Lemma 1, viz., the unique equilibrium outcome of the T-round game with initial default x^0 is $\phi^T(x^0)$. The more complicated statement here reflects the fact that, in settings with indifference, (i) both voters and the agenda setter can break ties differently across non-capricious equilibria and (ii) the agenda setter can break ties differently across rounds in a given non-capricious equilibrium. We prove part (a) by construction and part (b) using a backward-induction argument that leverages non-capriciousness; the proof is in Appendix B.2.2.

The primary import of Lemma 5 is that it implies bounds on the sets of outcomes and agenda setter payoffs across all non-capricious equilibria and initial defaults as the number of rounds becomes large.

Lemma 6. *For any collective choice problem $\mathcal{C}$, the following hold:*

(a) For every $x^0 \in X$ and $T \in \mathbb{N}$, we have

$$\mathcal{E} \subseteq \bigcup_{x^0 \in X} G_{T+1}(x^0) \subseteq \bigcup_{x^0 \in X} G_T(x^0).$$

(b) For every $\delta > 0$, there exists some $T_\delta \in \mathbb{N}$ such that:

$$\text{If } T \geq T_\delta, \text{ then } u_A(x) \geq \min_{y \in \mathcal{E}} u_A(y) - \delta \text{ for all } x \in \bigcup_{x^0 \in X} G_T(x^0).$$ [3]

Lemma 6(a) establishes that the set of Non-Capricious equilibrium outcomes—across all such equilibria and all initial defaults—converges monotonically downward to some set $\mathcal{G}_\infty^{\text{NC}} \supseteq \mathcal{E}$. Lemma 6(b) further shows that $\mathcal{G}_\infty^{\text{NC}} \subseteq \{x \in X : u_A(x) \geq \min_{y\in\mathcal{E}} u_A(y)\}$. Together, these facts imply that the agenda setter's minimal payoff across all non-capricious equilibria is precisely $\min_{y\in\mathcal{E}} u_A(y)$ in the $T \to \infty$ limit. We prove Lemma 6 in Appendix B.2.3 below.

Step 3: Main Argument for Theorem 3. Theorem 3(a) follows immediately from the existence claim in Lemma 5(a). We now use Lemma 6 to prove Theorem 3(b).

First, we show that manipulability is sufficient for approximate dictatorial power. Let $\mathcal{C}$ be a Manipulable collective choice problem. Then $\mathcal{E} = X_A^*$ and $\min_{y\in\mathcal{E}} u_A(y) = u_A^*$. Let $\delta > 0$ be given. Lemma 6(b) implies that there exists some $T_\delta \in \mathbb{N}$ such that the agenda setter's payoff is at least $u_A^* - \delta$ in every Non-Capricious equilibrium of any game with $T \geq T_\delta$ rounds, regardless of the initial default.

Next, we show that manipulability is necessary for approximate dictatorial power. Let $\mathcal{C}$ be a collective choice problem that is not Manipulable. Then there exists some $x \in \mathcal{E}\backslash X_A^*$ and $\delta > 0$ such that $u_A(x) < u_A^* - \delta$. Because $x \in \mathcal{E}$, Lemma 6(a) implies that, given any $T \in \mathbb{N}$, there exists an initial default $x^0 \in X$ and Non-Capricious equilibrium $\sigma \in \Sigma^{\text{NC}}(x^0, T)$ such that the outcome is $g_T^\sigma(x^0) = x$; in fact, Lemma 4(a) and Lemma 5 together imply

[3] Note that $\min_{y\in\mathcal{E}} u_A(y)$ is well-defined because u_A is continuous by assumption and $\mathcal{E}$ is closed, as the definition of improvability (Definition 1) and continuity of players' preferences imply that the set $X\backslash\mathcal{E}$ of improvable policies is open.

we can always pick the initial default to be $x^0 = x$. Thus, the agenda setter does not have approximate dictatorial power. $\square$

B.2.2 Proof of Lemma 5

Throughout this section, we take $x^0 \in X$ and $T \in \mathbb{N}$ as given and consider the T-round game with initial default x^0. For any round $t \in \{1, \ldots, T\}$ and prevailing default $x^{t-1} \in X$, let $\mathcal{H}^t(x^{t-1})$ denote the set of round-t histories consistent with this default. For any strategy profile σ that satisfies Definition 5(a), recall that $g^\sigma_T(x^0) \in X$ is the induced outcome starting from the initial history; correspondingly, for each $t \in \{2, \ldots, T\}$ and $x^{t-1} \in X$, let $g^\sigma_{T,x^0}(x^{t-1} \mid t) \in X$ denote the induced continuation outcome if x^{t-1} is the prevailing default in round t. To simplify some statements, we also extend this notation to the final round by letting $g^\sigma_{T,x^0}(x^T \mid T+1) := x^T$. Finally, for $t \in \{2, \ldots, T\}$, let $G^\sigma_{T,x^0}(t) := \bigcup_{x^{t-1} \in X} \left\{ g^\sigma_{T,x^0}(x^{t-1} \mid t) \right\}$ denote the set of continuation outcomes arising across all round-t subgames; to ease notation, we also let $G^\sigma_{T,x^0}(1) := \{g^\sigma_T(x^0)\}$.

Proof of Part (a). Let a selection $\hat{\phi}(\cdot) \in \Phi^{\text{or}}(\cdot)$ be given. We construct a pure Markovian strategy profile σ in the T-round game with initial default x^0 as follows:

- The agenda setter always proposes $\hat{\phi}(x)$ when the prevailing default is x.
- Each voter $i \in N$ votes to approve a proposal y in round t when the prevailing default is x^{t-1} if and only if either
  (i) $\hat{\phi}^{T-t}(y) \succ_i \hat{\phi}^{T-t}(x^{t-1})$, or
  (ii) $\hat{\phi}^{T-t}(y) \sim_i \hat{\phi}^{T-t}(x^{t-1})$ and $\hat{\phi}^{T-t}(y) = \hat{\phi}^{T-t+1}(x^{t-1})$.

We claim that σ is a non-capricious equilibrium.

First, observe that σ satisfies Definition 5(a) by construction; it induces the desired outcome $g^\sigma_T(x^0) = \hat{\phi}^T(x^0)$ and the continuation outcomes $g^\sigma_{T,x^0}(x^{t-1} \mid t) = \hat{\phi}^{T-t+1}(x^{t-1})$. Second, observe that σ satisfies Definition 5(b) also by construction: at any round-t history $h^t \in \mathcal{H}^t(x^{t-1})$, each voter $i \in N$ votes to approve a proposal y if and only if either

(i*) voter i strictly prefers $g^\sigma_{T,x^0}(y \mid t+1)$, the continuation outcome from approval of y, over $g^\sigma_{T,x^0}(x^{t-1} \mid t+1)$, the continuation outcome from rejection of y; or

(ii*) voter i is indifferent between these continuation outcomes and $g^\sigma_{T,x^0}(y \mid t+1) = \hat{\phi}(g^\sigma_{T,x^0}(x^{t-1} \mid t+1))$.

Definition 5(b) is satisfied because the tie-breaking rule in (ii*) depends only on the continuation outcomes conditional on approval and rejection.

Finally, we claim that σ is an equilibrium, and hence satisfies Definition 5. Clearly, no voter has a strictly profitable deviation, so it suffices to consider the agenda setter's incentives.

Let $x^{t-1} \in X$ and a round-t history $h^t \in \mathcal{H}^t(x^{t-1})$ be given and let $\omega := g^{\sigma}_{T,x^0}(x^{t-1}|t+1)$. By construction, a proposal y passes if and only if $g^{\sigma}_{T,x^0}(y \mid t+1) \in M^{\text{s}}(\omega) \cup \{\hat{\phi}(\omega)\}$. Thus, the agenda setter can induce all and only continuation outcomes $z \in M^{\text{s}}(\omega) \cup \{\hat{\phi}(\omega), \omega\}$, where ω is achieved by proposing any y that does not pass. Because $\hat{\phi}(\omega) \in \Phi^{\text{or}}(\omega)$ implies that $\hat{\phi}(\omega)$ is optimal for the agenda setter within $M^{\text{as}}(\omega) \supseteq M^{\text{s}}(\omega) \cup \{\hat{\phi}(\omega), \omega\}$, it follows that any proposal y for which $g^{\sigma}_{T,x^0}(y \mid t+1) = \hat{\phi}(\omega)$ is a best response for the agenda setter. Therefore, observing that $g^{\sigma}_{T,x^0}(\hat{\phi}(x^{t-1}) \mid t+1) = \hat{\phi}^{T-t+1}(x^{t-1}) = \hat{\phi}(\omega)$ completes the proof.

Proof of Part (b). Let a non-capricious equilibrium $\sigma \in \Sigma^{\text{NC}}(x^0, T)$ be given. We establish the existence of the desired collection $\{\hat{\phi}_t\}_{t=1}^T$ of selections $\hat{\phi}_t(\cdot) \in \Phi^{\text{or}}(\cdot)$ through a series of claims. The first claim records useful properties of continuation play and outcomes in the final round, $t = T$.

Claim 1. *There exists a selection $\hat{\phi}_T(\cdot) \in \Phi^{or}(\cdot)$ and an acceptance correspondence $M^{\sigma} : X \rightrightarrows X$ with the following properties:*

(a) *For every $x \in X$, $M^s(x) \cup \{\hat{\phi}_T(x), x\} \subseteq M^{\sigma}(x) \subseteq M^w(x)$.*

(b) *For every $x^{T-1} \in X$ and round-T history $h^T \in \mathcal{H}^T(x^{T-1})$, a proposal y such that $y \neq x^{T-1}$ is accepted if and only if $y \in M^{\sigma}(x^{T-1})$.*

(c) *For every $x^{T-1} \in X$, $\hat{\phi}_T(x^{T-1}) \in \arg\max_{z \in M^{\sigma}(x^{T-1})} u_A(z)$.*

(d) *For every $x^{T-1} \in X$, $g^{\sigma}_{T,x^0}(x^{T-1} \mid T) = \hat{\phi}_T(x^{T-1})$.*

Proof. In the final round T, for any proposal y and prevailing default x^{T-1}, acceptance of the proposal leads to continuation outcome y and rejection leads to x^{T-1}. Because σ satisfies Definition 5(b), there exists a correspondence $M^{\sigma} : X \rightrightarrows X$ such that, for every default $x^{T-1} \in X$ and history $h^T \in \mathcal{H}^T(x^{T-1})$, a proposal $y \neq x^{T-1}$ is accepted if and only if $y \in M^{\sigma}(x^{T-1})$. This establishes part (b). Also note that we may include $x^{T-1} \in M^{\sigma}(x^{T-1})$ for all $x^{T-1} \in X$ (as asserted in part (a)) without loss of generality, as both passage and rejection of a proposal $y = x^{T-1}$ leads to continuation outcome x^{T-1} at every history in $\mathcal{H}^T(x^{T-1})$, and part (b) only concerns proposals $y \neq x^{T-1}$.

Let $x^{T-1} \in X$ and $h^T \in \mathcal{H}^T(x^{T-1})$ be given. The fact that the continuation of σ in this subgame is an equilibrium thereof implies that M^{σ} satisfies $M^{\text{s}}(\cdot) \subseteq M^{\sigma}(\cdot) \subseteq M^{\text{w}}(\cdot)$ and the continuation outcome, call it $y(h^T)$, satisfies $y(h^T) \in \arg\max_{z \in M^{\sigma}(x^{T-1})} u_A(z)$. Moreover, because this continuation equilibrium is Non-Capricious, Lemma 4(b) implies that $y(h^T) \in \Phi^{\text{or}}(x^{T-1}) \cap M^{\sigma}(x^{T-1})$. Finally, because σ satisfies Definition 5(a), there exists some $\hat{\phi}(x^{T-1}) \in \Phi^{\text{or}}(x^{T-1}) \cap M^{\sigma}(x^{T-1})$ such that $y(h^T) = \hat{\phi}(x^{T-1})$ for all $h^T \in \mathcal{H}^T(x^{T-1})$ and $x^{T-1} \in X$. This establishes parts (a), (c), and (d). □

The next claim uses the Non-Capricious refinement to show that the majority acceptance correspondence from Claim 1 also characterizes voter behavior in all rounds, and records a useful implication of this fact.

Claim 2. *For every $1 \leq t \leq T$, default $x^{t-1} \in X$, and round-t history $h^t \in \mathcal{H}^t(x^{t-1})$, the following hold:*

(a) *A proposal y such that $g^\sigma_{T,x^0}(y \mid t+1) \neq g^\sigma_{T,x^0}(x^{t-1} \mid t+1)$ is accepted at h^t if and only if $g^\sigma_{T,x^0}(y \mid t+1) \in M^\sigma\left(g^\sigma_{T,x^0}(x^{t-1} \mid t+1)\right)$.*

(b) *The continuation outcome at h^t satisfies*

$$g^\sigma_{T,x^0}(x^{t-1} \mid t) \in \arg\max\left\{u_A(z) : z \in M^\sigma\left(g^\sigma_{T,x^0}(x^{t-1} \mid t+1)\right) \bigcap G^\sigma_{T,x^0}(t+1)\right\}.$$

Proof. Part (a) follows directly from Claim 1(b) and the fact that σ satisfies Definition 5. Part (b) follows directly from part (a) and the fact that σ satisfies Definition 5. □

The next claim records the elementary observation that any continuation outcome of a round-t subgame must also be the continuation outcome of some round-$(t+1)$ subgame.

Claim 3. *For every $1 \leq t \leq T-1$, we have $G^\sigma_{T,x^0}(t) \subseteq G^\sigma_{T,x^0}(t+1)$.*

Proof. Let the round $t \in \{1, \ldots, T-1\}$, default $x^{t-1} \in X$, and history $h^t \in \mathcal{H}^t(x^{t-1})$ be given. By definition, the continuation outcome at h^t is $g^\sigma_{T,x^0}(x^{t-1} \mid t)$. If σ specifies that some proposal $a^\sigma(h^t) \in X$ be made and passed with positive probability at h^t, then by construction we have $g^\sigma_{T,x^0}(x^{t-1} \mid t) = g^\sigma_{T,x^0}(a^\sigma(h^t) \mid t+1)$. If σ specifies that no proposals pass with positive probability at h^t, then by construction we have $g^\sigma_{T,x^0}(x^{t-1} \mid t) = g^\sigma_{T,x^0}(x^{t-1} \mid t+1)$. The claim now follows immediately from the definition of $G^\sigma_{T,x^0}(t)$ and $G^\sigma_{T,x^0}(t+1)$. □

The final claim uses Claims 1 to 3 to characterize outcomes at every history. Lemma 5(b) is directly implied by this claim.

Claim 4. *There exists a collection $\{\hat{\phi}_t\}_{t=1}^T$ of selections $\hat{\phi}_t(\cdot) \in \Phi^{or}(\cdot) \cap M^\sigma(\cdot)$ such that the following hold:*

(a) *For every $2 \leq t \leq T$, the continuation outcomes satisfy the following:*

$$\text{For all } x^{t-1} \in X, \;\; g^\sigma_{T,x^0}(x^{t-1} \mid t) = \left[\hat{\phi}_t \circ \hat{\phi}_{t+1} \circ \cdots \circ \hat{\phi}_T\right](x^{t-1}). \tag{13}$$

Analogously, the equilibrium outcome of the game is $g^\sigma_T(x^0) = \left[\hat{\phi}_1 \circ \hat{\phi}_2 \circ \cdots \circ \hat{\phi}_T\right](x^0)$.

(b) *For every $x \in X$ and $1 \leq t \leq T$, we have $\hat{\phi}_t(x) \sim_A \hat{\phi}_T(x)$.*

Proof. We prove (a) by backward induction. Let $\hat{\phi}_T(\cdot) \in \Phi^{\text{or}}(\cdot) \cap M^\sigma(\cdot)$ be as defined in Claim 1. Claim 1(d) then establishes the base ($t = T$) case of (13). By letting $z := \hat{\phi}_T(x^{T-1})$ and $z' := \hat{\phi}_T(z)$ for any given $x^{T-1} \in X$, Claim 1(d) also establishes the base ($t = T$) case of the following property:[4]

$$\text{If } z \in G^\sigma_{T,x^0}(t), \text{ then } \exists\, z' \in G^\sigma_{T,x^0}(t) \cap M^\sigma(z) \cap \Phi^{\text{or}}(z) \text{ such that } z' \succsim_A \hat{\phi}_T(z). \tag{14}$$

For the inductive step, suppose for a given $\tau \in \{2, \dots, T-1\}$ that (i) the selections $\{\hat{\phi}_s\}_{s=t}^T$ of Φ^{or} satisfy (13) for $t = \tau + 1$ and (ii) (14) holds for $t = \tau + 1$.

We first assert that there exists a selection $\hat{\phi}_\tau(\cdot) \in \Phi^{\text{or}}(\cdot)$ satisfying (13) for $t = \tau$. Let $x^{\tau-1} \in X$ and $h^\tau \in \mathcal{H}^\tau(x^{\tau-1})$ be given. Let $z := g^\sigma_{T,x^0}(x^{\tau-1} \mid \tau+1) \in G^\sigma_{T,x^0}(\tau+1)$ denote the continuation outcome if $x^{\tau-1}$ remains the default in the next round, $t = \tau+1$. By the inductive hypothesis that (14) holds for $t = \tau + 1$ and Claim 2(b), the continuation outcome at h^τ, which is $g^\sigma_{T,x^0}(x^{\tau-1} \mid \tau)$, must satisfy $g^\sigma_{T,x^0}(x^{\tau-1} \mid \tau) = \hat{\phi}_\tau(z)$ for some $\hat{\phi}_\tau(z) \in \Phi^{\text{or}}(z) \cap M^\sigma(z)$ such that $\hat{\phi}_\tau(z) \succsim_A \hat{\phi}_T(z)$. Now, repeating this logic across all round-τ histories delivers, for all $x \in G^\sigma_{T,x^0}(\tau + 1)$, the existence of some $\hat{\phi}_\tau(x) \in \Phi^{\text{or}}(x) \cap M^\sigma(x)$ such that (13) holds for $t = \tau$ and $\hat{\phi}_\tau(x) \succsim_A \hat{\phi}_T(x)$. Since no policy in $X \backslash G^\sigma_{T,x^0}(\tau + 1)$ can be induced as a continuation outcome by any proposal at any round-τ history, we may arbitrarily assign $\hat{\phi}_\tau(x) := \hat{\phi}_T(x) \in \Phi^{\text{or}}(x) \cap M^\sigma(x)$ for each $x \in X \backslash G^\sigma_{T,x^0}(\tau + 1)$. This results in the desired selection $\hat{\phi}_\tau(\cdot) \in \Phi^{\text{or}}(\cdot)$, completing the proof of the assertion.

Next, we assert that (14) holds for $t = \tau$. Let $z \in G^\sigma_{T,x^0}(\tau)$ be given. Claim 3 implies that $z \in G^\sigma_{T,x^0}(\tau + 1)$, so that $z = g^\sigma_{T,x^0}(x \mid \tau + 1)$ for some round-$(\tau + 1)$ default $x \in X$. Let $z' := g^\sigma_{T,x^0}(x \mid \tau)$ denote the continuation outcome if x is the round-τ default. By the argument in the preceding paragraph, we have $z' = \hat{\phi}_\tau(z) \in M^\sigma(z) \cap \Phi^{\text{or}}(z)$ and thus $z' \succsim_A \hat{\phi}_T(z)$. As $z' \in G^\sigma_{T,x^0}(\tau)$ by construction, the assertion is proved.

This completes the inductive proof of part (a) for all rounds $t \in \{2, \dots, T\}$. Repeating the first inductive step above once more establishes it for round $t = 1$.

To prove part (b), note that $\hat{\phi}_t(x) \succsim_A \hat{\phi}_T(x)$ for all $1 \le t \le T$ and $x \in X$ by construction. Suppose, towards a contradiction, that there exists some $1 \le t \le T$ and $x \in X$ such that $\hat{\phi}_t(x) \succ_A \hat{\phi}_T(x)$. Then consider any round-T history $h^T(x)$ in which the default is $x^{T-1} = x$. Because $\hat{\phi}_t(x) \in M^\sigma(x)$ by construction, Claim 1(b) implies that the agenda setter has a strictly profitable deviation at h^T by proposing $\hat{\phi}_t(x)$ instead of $\hat{\phi}_T(x)$, contradicting that σ is an equilibrium. □

[4] Specifically, if $z \in G^\sigma_{T,x^0}(T)$ then, by definition, $z = g^\sigma_{T,x^0}(x^{T-1} \mid T)$ for some $x^{T-1} \in X$. Claim 1(d) then implies that $z = \hat{\phi}_T(x^{T-1})$. Analogously, Claim 1(d) implies that $z' := \hat{\phi}_T(z) = g^\sigma_{T,x^0}(z \mid T) \in G^\sigma_{T,x^0}(T)$. Hence, by construction, $z' \in G^\sigma_{T,x^0}(T) \cap M^\sigma(z) \cap \Phi^{\text{or}}(z)$ and $z' \sim_A \hat{\phi}_T(z)$, which yields the base ($t = T$) case of (14).

B.2.3 Proof of Lemma 6

In this section, we use the same notation introduced at the beginning of Appendix B.2.3. We prove each part of Lemma 6 in turn.

Proof of Part (a). We first show that every unimprovable policy is a Non-Capricious equilibrium outcome. Let $x \in \mathcal{E}$ be given. Lemma 4(a) implies that $x \in \Phi^{\text{or}}(x)$; hence, there exists a selection $\hat{\phi}(\cdot) \in \Phi^{\text{or}}(\cdot)$ such that $\hat{\phi}(x) = x$. Lemma 5(a) then implies that, for every $T \in \mathbb{N}$, there exists a Non-Capricious equilibrium $\sigma \in \Sigma^{\text{NC}}(x, T)$ with outcome $g_T^\sigma(x) = x$. Thus, $\mathcal{E} \subseteq \bigcup_{x^0 \in X} G_T(x^0)$ for every $T \in \mathbb{N}$.

Next, we show that the equilibrium outcome sets are decreasing in the number of rounds. Let $T \in \mathbb{N}$, $x^0 \in X$, and $\sigma \in \Sigma^{\text{NC}}(x^0, T)$ be given. By Claim 3 in Appendix B.2.2, we have $\{g_T^\sigma(x^0)\} = G_{T,x^0}^\sigma(1) \subseteq G_{T,x^0}^\sigma(2)$. As the continuation of σ at any round-2 history (of the T-round game with initial default x^0) is a Non-Capricious equilibrium in the corresponding $(T-1)$-round subgame, it follows that $G_{T,x^0}^\sigma(2) \subseteq \bigcup_{y^0 \in X} G_{T-1}(y^0)$. It follows that

$$\bigcup_{x^0 \in X} G_T(x^0) \subseteq \bigcup_{x^0 \in X} \bigcup_{\sigma \in \Sigma^{\text{NC}}(x^0, T)} G_{T,x^0}^\sigma(2) \subseteq \bigcup_{y^0 \in X} G_{T-1}(y^0),$$

which completes the proof.

Proof of Part (b). We begin by stating a useful variant of the *uniform improvement lemma* (Lemma 2) used in Appendix A.2 to prove Theorem 2. For each $\delta > 0$, let

$$\Upsilon_\delta := \left\{ x \in X : \min_{y \in \mathcal{E}} u_A(y) \geq u_A(x) + \delta \right\}$$

denote the set of policies that are δ-dominated for the agenda setter by *all* unimprovable policies $y \in \mathcal{E}$. Obviously, $\Upsilon_\delta \subseteq X \backslash \mathcal{E}$ for all $\delta > 0$. As in Appendix A.2, for each $x \in X$ and $\eta > 0$, define

$$\begin{aligned} Q(x, \eta) := \big\{ y \in X \,|u_A(y) \geq u_A(x) + \eta \text{ and} \\ \exists \text{ majority } S \subseteq N \text{ such that } u_i(y) \geq u_i(x) + \eta \ \ \forall i \in S \big\} \end{aligned}$$

to be the set of policies that lead to a utility improvement of at least η for some winning coalition. If $Q(x, \eta) \neq \emptyset$, then we say that x is η-improvable.

Lemma 7. *For any collective choice problem $\mathcal{C}$ and every $\delta > 0$, there exists $\eta_\delta > 0$ such that $Q(x, \eta_\delta) \neq \emptyset$ for all $x \in \Upsilon_\delta$.*

Proof. Let $\delta > 0$ be given. Suppose that $\Upsilon_\delta \neq \emptyset$, for otherwise the lemma is vacuously true. First, observe that Υ_δ is compact because u_A is continuous and X is compact. Second, observe that for each $x \in \Upsilon_\delta$ there exists some $\eta_x > 0$ such that $Q(x, \eta_x) \neq \emptyset$; this follows from the definition of improvability and the inclusion $\Upsilon_\delta \subseteq X \backslash \mathcal{E}$. Given these observations, the remainder of the proof is identical to the proof of Lemma 2 in Appendix A.2.[5] □

Now, let $\delta > 0$ be given and let $\eta_\delta > 0$ be as described in Lemma 7. By the definitions of $V_A^{\text{as}}(\cdot)$ and $\Phi^{\text{or}}(\cdot)$, we have

$$u_A(z) - u_A(y) \geq V_A^{\text{as}}(y) - u_A(y) \geq \eta_\delta \quad \text{for all } y \in \Upsilon_\delta \text{ and } z \in \Phi^{\text{or}}(y). \tag{15}$$

Let $\Delta := u_A^* - \min_{x \in X} u_A(x)$. We prove Lemma 6(b) by showing that

$$T \geq T_\delta := \lceil \Delta / \eta_\delta \rceil \quad \Longrightarrow \quad \bigcup_{x^0 \in X} G_T(x^0) \subseteq X \backslash \Upsilon_\delta. \tag{16}$$

Towards a contradiction, suppose that there exists a default $x^0 \in X$, number of rounds $T \geq T_\delta$, and Non-Capricious equilibrium $\sigma \in \Sigma^{\text{NC}}(x^0, T)$ such that $g_T^\sigma(x^0) \in \Upsilon_\delta$. By Claim 4(a), $g_T^\sigma(x^0) = \left[\hat{\phi}_1 \circ \cdots \circ \hat{\phi}_T\right](x^0)$ and $g_{T,x^0}^\sigma(x^0 \mid t) = \left[\hat{\phi}_t \circ \cdots \circ \hat{\phi}_T\right](x^0)$ for some selections $\hat{\phi}_t(\cdot) \in \Phi^{\text{or}}(\cdot)$. Let $g_{T,x^0}^\sigma(x^0 \mid 1) := g_T^\sigma(x^0)$. Note that $g_{T,x^0}^\sigma(x^0 \mid t) \in \Phi^{\text{or}}(g_{T,x^0}^\sigma(x^0 \mid t+1))$ for all $1 \leq t \leq T$. Then it follows that

$$\begin{aligned}
V_A^{\text{as}}\left(g_T^\sigma(x^0)\right) - u_A(x^0) &\geq u_A\left(g_T^\sigma(x^0)\right) - u_A(x^0) + \eta_\delta \\
&= \sum_{t=1}^{T} \left[u_A\left(g_{T,x^0}^\sigma(x^0 \mid t)\right) - u_A\left(g_{T,x^0}^\sigma(x^0 \mid t+1)\right)\right] + \eta_\delta \\
&\geq T_\delta \cdot \eta_\delta + \eta_\delta \\
&\geq \Delta + \eta_\delta,
\end{aligned}$$

where the first line is by (15), the second line is an identity, the third line is by the hypothesis that $T \geq T_\delta$ and another application of (15) to each term in the sum (noting that $g_T^\sigma(x^0) \in \Upsilon_\delta$ implies that $g_{T,x^0}^\sigma(x^0 \mid t) \in \Upsilon_\delta$ for all $1 \leq t \leq T+1$), and the final line is by definition of T_δ. However, given that $\eta_\delta > 0$, this inequality contradicts the definition of Δ. We conclude that (16) holds, as desired.

B.3 Proof of Theorem 4

Our argument proceeds in several steps. First, in 3 dimensions, we establish a connection between non-coplanarity of utility gradients and improvability; as the reader will see, this

[5] The only difference is that here we appeal to the definition of improvability and the inclusion $\Upsilon_\delta \subseteq X \backslash \mathcal{E}$—rather than the manipulability of $\mathcal{C}$—to establish the second observation above.

argument applies for general payoff functions. We use this step to prove Theorem 4(a) for $d \geq 3$, by doing the appropriate reduction to 3 dimensions and noting that non-coplanarity of ideal points with Euclidean preferences (Non-Coplanarity) implies that of utility gradients. Finally, we prove Theorem 4(b) directly.

Step 1: A General Improvability Lemma for 3 Dimensions. We establish here, for general payoff functions, that if utility gradients are non-coplanar at policy $x \neq x_A^*$, policy x must be improvable.

Lemma 8. *Suppose $X = \mathbb{R}^3$ and each player $i \in N \cup \{A\}$ has a strictly quasi-concave and continuously differentiable utility $v_i : X \to \mathbb{R}$ with unique maximizer x_i^*. Then any policy $x \neq x_A^*$ satisfying $\nabla v_A(x) \neq \mathbf{0}$ is improvable if*

$$\text{no four players' gradients, } \nabla v_1(x), \ldots, \nabla v_n(x), \nabla v_A(x) \in \mathbb{R}^3, \text{ are coplanar.} \tag{17}$$

Lemma 8 formalizes, for general preferences, the argument from our proof sketch for Euclidean preferences in the main text.[6]

Proof of Lemma 8. Let the profiles $(v_i)_{i=1,\ldots,n,A}$ and $(x_i^*)_{i=1,\ldots,n,A}$ be as described above, and consider a policy $x \in \mathbb{R}^3 \backslash \{x_A^*\}$ that satisfies $\nabla v_A(x) \neq \mathbf{0}$ and (17) . Denote the plane tangent to the agenda setter's indifference surface at x by $S := \{y \in \mathbb{R}^3 : (y - x) \cdot \nabla v_A(x) = 0\}$. The tangent space of S is denoted by $\mathcal{T}(S) := \{z \in \mathbb{R}^3 : z \cdot \nabla v_A(x) = 0\}$ and the orthogonal complement of S by $S^\perp := \{y \in \mathbb{R}^3 : y \cdot z = 0 \ \forall z \in \mathcal{T}(S)\}$. For each voter $i \in N$, denote the orthogonal projection of $\nabla v_i(x)$ onto S by $\nabla_S v_i(x) := \nabla v_i(x) - \left(\frac{\nabla v_i(x) \cdot \nabla v_A(x)}{\|\nabla v_A(x)\|^2}\right) \nabla v_A(x)$. By construction, $\nabla_S v_i(x) \in \mathcal{T}(S)$ and $\nabla_S v_i(x) \cdot y = \nabla v_i(x) \cdot y$ for all $y \in \mathcal{T}(S)$.

We now establish Lemma 8 through a sequence of claims, which parallel the geometric sketch for 3-dimensional Euclidean preferences given in Section 4.1. The first claim records useful implications of (17) for the voters' projected gradients.

Claim 5. *The following hold:*

(a) *There exists at most one voter $i \in N$ for whom $\nabla_S v_i(x) = \mathbf{0}$.*

(b) *Given any voter $i \in N$ for whom $\nabla_S v_i(x) \neq \mathbf{0}$, there exists at most one other voter $j \in N \backslash \{i\}$ with a parallel projected gradient, viz., such that $\nabla_S v_j(x) = \alpha \cdot \nabla_S v_i(x)$ for some $\alpha \in \mathbb{R}$.*[7]

[6] For Euclidean preferences, player i's gradient at x is $\nabla v_i(x) = x_i^* - x$, where x_i^* is i's ideal point. Plainly, any policy $x \neq x_A^*$ satisfies $\nabla v_A(x) \neq \mathbf{0}$. Furthermore, for any four players $\{i, j, k, \ell\}$ and any policy $x \in \mathbb{R}^3$, $\{x_i^*, x_j^*, x_k^*, x_\ell^*\}$ are coplanar if and only if $\{x_i^* - x, x_j^* - x, x_k^* - x, x_\ell^* - x\}$ are coplanar. Therefore, Non-Coplanarity implies that (17) holds for *all* $x \in \mathbb{R}^3$; conversely, if (17) holds for *some* $x \in \mathbb{R}^3$, then Non-Coplanarity holds.

[7] Equivalently, there exists at most one voter $j \in N \backslash \{i\}$ such that the projected gradients $\nabla_S v_i(x)$ and $\nabla_S v_j(x)$ are collinear with $\nabla_S v_A(x) = \mathbf{0}$.

Claim 5 formalizes, for general preferences, the analogs of (i) and (ii) from Step 1 of our proof sketch for Euclidean preferences in the main text.[8]

Proof of Claim 5. For part (a), suppose that $\nabla_S v_i(x) = \nabla_S v_j(x) = \mathbf{0}$ for two distinct voters $i, j \in N$. By definition, $\nabla_S v_i(x) = \mathbf{0}$ (resp., $\nabla_S v_j(x) = \mathbf{0}$) if and only if $\nabla v_i(x) \in S^\perp$ (resp., $\nabla v_j(x) \in S^\perp$). Because $S^\perp$ has dimension 1 and contains $\nabla v_A(x)$, it follows that the gradients $\{\nabla v_i(x), \nabla v_j(x), \nabla v_A(x)\}$ are collinear. Thus, for any third voter $k \in N$, the gradients $\{\nabla v_i(x), \nabla v_j(x), \nabla v_A(x), \nabla v_k(x)\}$ are coplanar. (17) is then violated. By contraposition, (17) implies that (a) holds.

For part (b), consider a voter $i \in N$ for whom $\nabla_S v_i(x) \neq \mathbf{0}$. Suppose that there exist two distinct voters $j, k \in N \backslash \{i\}$ such that the projected gradients $\{\nabla_S v_i(x), \nabla_S v_j(x), \nabla_S v_k(x)\}$ are parallel. Observe that these vectors are also trivially parallel to $\nabla_S v_A(x) = \mathbf{0}$. Hence, the space $\operatorname{span}(\{\nabla_S v_i(x), \nabla_S v_j(x), \nabla_S v_k(x), \nabla_S v_A(x)\})$ has dimension 1. Note that $\nabla v_\nu(x) = \nabla_S v_\nu(x) + \nabla_{S^\perp} v_\nu(x)$ for all $\nu \in \{i, j, k, A\}$ by definition of orthogonal projection. Because each $\nabla_{S^\perp} v_\nu(x) \in S^\perp$ and $S^\perp$ has dimension 1 by construction, it follows that the space $\operatorname{span}(\{\nabla v_i(x), \nabla v_j(x), \nabla v_k(x), \nabla v_A(x)\})$ has dimension 2, implying that (17) is violated. □

The next claim uses Claim 5 to establish the existence of an alternative policy in S that a majority of voters strictly prefer to x.

Claim 6. *There exists some $y \in S$ such that $y \succ_M x$.*

The argument mirrors that from Step 1 and the left-hand panel of Figure 4 in Section 4.1: we (i) fix some voter $i \in N$ whose projected gradient $\nabla_S v_i(x)$ is nonzero and therefore defines a line in S that contains x, (ii) partition the other voters into sets according to whether their projected gradients point "above" or "below" that line, and (iii) construct a new policy $y \in S$ that strictly benefits voter i and all voters on one side of the line. For Euclidean preferences, part (ii) here is equivalent to partitioning voters based on their constrained ideal points y_j^* lying "above" or "below" the line.

Proof of Claim 6. Consider a voter i for whom $\nabla_S v_i(x) \neq \mathbf{0}$; such a voter exists by Claim 5(a). We denote the set of other voters whose S-projected gradients at policy x are parallel to i's by $C_i := \{j \in N \backslash \{i\} : \nabla_S v_j(x) = \alpha \cdot \nabla_S v_i(x) \; \exists \alpha \in \mathbb{R}\}$. We then define $N' := N \backslash C_i$. Observe that $i \in N'$ by construction and that $|N'| \geq n - 1$ by Claim 5(b).

[8] For Euclidean preferences, player i's S-projected gradient at x is $\nabla_S v_i(x) = y_i^* - x$, where y_i^* is i's constrained ideal point in S.

Now, let $\omega \in \mathcal{T}(S)\backslash\{\mathbf{0}\}$ satisfying $\omega \cdot \nabla_S v_i(x) = 0$ be given. Define the following sets of voters:

$$N'_+ := \{j \in N' : \nabla_S v_j(x) \cdot \omega > 0\} \quad \text{and} \quad N'_- := \{j \in N' : \nabla_S v_j(x) \cdot \omega < 0\}.$$

By construction, $N'_+ \cap \{i\} = N'_- \cap \{i\} = \emptyset$ and $N' = N'_+ \cup N'_- \cup \{i\}$, viz., $\{N'_+, N'_-, \{i\}\}$ forms a partition of N'. It follows that $|N'_+| + |N'_-| \geq n - 2$, which in turn implies that $\max\{|N'_+|, |N'_-|\} \geq \frac{n-1}{2}$ because $n - 2$ is an odd number. We suppose, without loss of generality, that $|N'_+| \geq \frac{n-1}{2}$. Therefore, we have

$$|N'_+ \cup \{i\}| \geq \frac{n+1}{2}. \tag{18}$$

We assert that there exists some $\rho \in \mathbb{R}^3$ such that

$$\rho \in \mathcal{T}(S) \text{ and } \nabla_S v_j(x) \cdot \rho > 0 \text{ for all } j \in N'_+ \cup \{i\}. \tag{19}$$

To this end, define the sequence $\{\rho^k\} \subset \mathbb{R}^3$ by $\rho^k := \frac{1}{k}\nabla_S v_i(x) + \frac{k-1}{k}\omega$. It is clear that $\rho^k \in \mathcal{T}(S)$ for all $k \in \mathbb{N}$, as $\mathcal{T}(S)$ is a convex set containing both $\nabla_S v_i(x)$ and ω by construction. It is also clear that $\nabla_S v_i(x) \cdot \rho^k = \|\nabla_S v_i(x)\|^2/k > 0$ for all $k \in \mathbb{N}$ by construction. So, let $j \in N'_+$ be given. As $\nabla_S v_j(x) \cdot \omega > 0$ by construction, there exists some $K_j \in \mathbb{N}$ such that $\nabla_S v_j(x) \cdot \rho^k > 0$ for all $k \geq K_j$. Defining $K := \max_{j \in N'_+} K_j$ and letting $\rho := \rho^k$ for any $k \geq K$ then establishes (19), as desired.

We now use (18) and (19) to prove the claim. Let $\rho \in \mathbb{R}^3$ satisfy (19). For each voter $j \in N'_+ \cup \{i\}$, we have that

$$v_j(x + \epsilon\rho) = v_j(x) + \epsilon\nabla v_j(x) \cdot \rho + o(\epsilon) = v_j(x) + \epsilon\nabla_S v_j(x) \cdot \rho + o(\epsilon),$$

where the first equality is by Taylor's Theorem and the second holds because, by the definition of the S-projected gradient $\nabla_S v_j(x)$, we have $\nabla v_j(x) \cdot \rho' = \nabla_S v_j(x) \cdot \rho'$ for all $\rho' \in \mathcal{T}(S)$. (19) then implies that there exists some $\epsilon > 0$ such that $v_j(x + \epsilon\rho) > v_j(x)$ for all $j \in N'_+ \cup \{i\}$. Letting $y := x + \epsilon\rho$, (18) then implies that $y \succ_M x$. As $\rho \in \mathcal{T}(S)$, it follows that $y \in S$. □

The final claim establishes that any policy $y \in S$ for which $y \succ_M x$ can be perturbed to some $z \notin S$ such that both $z \succ_M x$ and $z \succ_A x$; this claim formalizes the argument sketched in Step 2 and the right-hand panel of Figure 4.

Claim 7. *For any $y \in S$ such that $y \succ_M x$, there exists $z \notin S$ such that $z \succ_M x$ and $z \succ_A x$.*

Proof. As $y \succ_M x$ and voters' preferences are continuous, there exists an $\epsilon > 0$ such that the policy $\zeta := y + \epsilon \nabla v_A(x)$ satisfies $\zeta \succ_M x$. For each $\beta \in (0,1)$, define $z(\beta) := \beta\zeta + (1-\beta)x$. The strict convexity of voters' preferences then implies that $z(\beta) \succ_M x$ for all $\beta \in (0,1)$.

We assert that there exists some $\overline{\beta} \in (0,1)$ such that $z(\beta) \succ_A x$ for all $\beta \in (0, \overline{\beta})$. Let $\rho := \zeta - x$. As v_A is continuously differentiable, its directional derivative at policy x in direction ρ is given by $\nabla v_A(x) \cdot \rho$. We have the following:

$$\nabla v_A(x) \cdot \rho = \nabla v_A(x) \cdot (y + \epsilon \nabla v_A(x) - x) = \epsilon \|\nabla v_A(x)\|^2 > 0$$

where the first equality is an identity, and the second follows from rearranging terms and noting that $\nabla v_A(x) \neq \mathbf{0}$ and $\nabla v_A(x) \perp (y - x)$. Thus, we have $\nabla v_A(x) \cdot \rho > 0$. Now, because $z(\beta) - x = \beta\rho$ by construction, Taylor's Theorem implies that $v_A(z(\beta)) = v_A(x) + \beta \nabla v_A(x) \cdot \rho + o(\beta)$. It follows that $v_A(z(\beta)) > v_A(x)$ for all sufficiently small $\beta > 0$, as desired.

Now let $z := z(\beta)$ for any $\beta \in (0, \overline{\beta})$. It follows that $z \succ_A x$ and $z \succ_M x$ by construction and $z \notin S$ because $\nabla v_A(x) \neq \mathbf{0}$ is normal to S. □

Claims 6 and 7 together complete the proof of Lemma 8. □

Step 2: Proof of Theorem 4(a). We now consider the case of Euclidean preferences: $d \geq 3$, $X = \mathbb{R}^d$, player i has utility function $u_i(x) = -\frac{1}{2}\|x - x_i^*\|^2$ for each $i \in N \cup \{A\}$. Suppose that the ideal point profile $(x_i^*)_{i=1,\ldots,n,A} \in \mathbb{R}^{d(n+1)}$ satisfies Non-Coplanarity. Let an arbitrary $x \neq x_A^*$ be given; we show below that x is improvable.

If $d = 3$, the result follows immediately from Lemma 8 by observing that $\nabla u_i(x) = x_i^* - x$, so that Non-Coplanarity directly implies (17). So suppose that $d > 3$. In this case, we may indirectly apply Lemma 8 by restricting attention to a suitable 3-dimensional subspace of $\mathbb{R}^d$. Let $a, b, c \in \{1, \ldots, d\}$ denote 3 distinct policy dimensions for which the projections $[x]_{abc}$ and $[x_A^*]_{abc}$ satisfy $[x]_{abc} \neq [x_A^*]_{abc}$. Let $[x]_{-abc} \in \mathbb{R}^{d-3}$ denote the $(d-3)$-dimensional projection of x corresponding to deletion of the indices a, b, c (so that x is given by the concatenation of $[x]_{abc}$ and $[x]_{-abc}$). Define $X([x]_{-abc}) := \{y \in \mathbb{R}^d : [y]_{-abc} = [x]_{-abc}\}$ to be the set of policies $y \in \mathbb{R}^d$ that differ from $x \in \mathbb{R}^d$ only in dimensions a, b, c. Observe that $X([x]_{-abc})$ is a 3-dimensional affine subspace of $\mathbb{R}^d$ by construction; with a slight abuse of notation, we identify it with $\mathbb{R}^3$ and identify a generic element y with its projection $[y]_{abc} \in \mathbb{R}^3$. Finally, for each player $i \in N \cup \{A\}$, we define the utility function $v_i : \mathbb{R}^3 \to \mathbb{R}$ by $v_i(\cdot) := u_i(\cdot, [x]_{-abc})$, viz., v_i is the restriction of u_i to $X([x]_{-abc})$.

We now apply Lemma 8 to the utility profile $(v_i)_{i=1,\ldots,n,A}$, which represents 3-dimensional Euclidean preferences with the ideal point profile $([x_i^*]_{abc})_{i=1,\ldots,n,A} \in \mathbb{R}^{3(n+1)}$. Observe that $\nabla v_i([x]_{abc}) = [x_i^*]_{abc} - [x]_{abc} \in \mathbb{R}^3$, so that Non-Coplanarity of the d-dimensional ideal points implies that these 3-dimensional gradients satisfy (17). We thus conclude from Lemma 8 that there exists some $[y]_{abc} \in \mathbb{R}^3$ such that

$$v_A([y]_{abc}) > v_A([x]_{abc}) \quad \text{and} \quad |\{i \in N : v_i([y]_{abc}) > v_i([x]_{abc})\}| \geq \frac{n+1}{2}. \tag{20}$$

To conclude the proof, we let $y \in X([x]_{-abc}) \subseteq \mathbb{R}^d$ denote the concatenation of $[y]_{abc}$ and $[x]_{-abc}$, viz., $y := ([y]_{abc}, [x]_{-abc})$. By definition of the v_i functions, (20) implies that

$$u_A(y) > u_A(x) \quad \text{and} \quad |\{i \in N : u_i(y) > u_i(y)\}| \geq \frac{n+1}{2},$$

which is equivalent to $y \succ_A x$ and $y \succ_M x$. It follows that x is improvable, as desired. □

Step 3: Proof of Theorem 4(b). Let $d \geq 3$ be given. For any 3 distinct policy dimensions $a, b, c \in \{1, \ldots, d\}$ and any 4 distinct players $i, j, k, \ell \in \{1, \ldots, n, A\}$, we define the set $C^{(ijk\ell)}_{[abc]} \subseteq \mathbb{R}^{d(n+1)}$ by

$$C^{(ijk\ell)}_{[abc]} := \left\{ (x^*_\nu)_{\nu=1,\ldots,n,A} \in \mathbb{R}^{d(n+1)} : [x^*_i]_{abc}, [x^*_j]_{abc}, [x^*_k]_{abc}, \text{ and } [x^*_\ell]_{abc} \text{ are coplanar in } \mathbb{R}^3 \right\}.$$

In words, $C^{(ijk\ell)}_{[abc]}$ collects all profiles of ideal points for which Non-Coplanarity is violated (at least) for players i, j, k, ℓ in the subspace spanned by dimensions a, b, c. Observe that, by Definition 6, taking the union over all such a, b, c and i, j, k, ℓ yields exactly the set of ideal point profiles that violate Non-Coplanarity. That is, the following holds:

$$\begin{aligned} C &:= \left\{ (x^*_i)_{i=1,\ldots,n,A} \in \mathbb{R}^{d(n+1)} : (x^*_i)_{i=1,\ldots,n,A} \text{ violates Non-Coplanarity} \right\} \\ &= \bigcup_{a,b,c \in \{1,\ldots,d\}} \bigcup_{i,j,k,\ell \in N \cup \{A\}} C^{(ijk\ell)}_{[abc]}. \end{aligned} \tag{21}$$

We show that C is closed and has zero Lebesgue measure by showing that each $C^{(ijk\ell)}_{[abc]}$ also has these properties (because the union in (21) is finite). To this end, we first claim that

$$Z := \left\{ (x_i)_{i=1}^4 \in \mathbb{R}^{3\times 4} : x_i \in \mathbb{R}^3 \ \forall i = 1,2,3,4 \text{ and } x_1, x_2, x_3, x_4 \text{ are coplanar in } \mathbb{R}^3 \right\} \tag{22}$$

is closed and has zero Lebesgue measure. To see why, define $f : \mathbb{R}^{3\times 4} \to \mathbb{R}$ by $f(x_1, x_2, x_3, x_4) := [(x_2 - x_1) \times (x_3 - x_1)] \cdot (x_4 - x_1)$, where $y \times z \in \mathbb{R}^3$ denotes the cross product between vectors $y, z \in \mathbb{R}^3$. By construction, $\{x_1, x_2, x_3, x_4\}$ are coplanar if and only if $f(x_1, x_2, x_3, x_4) = 0$; therefore, $Z = \{(x_i)_{i=1}^4 \in \mathbb{R}^{3\times 4} : f(x_1, x_2, x_3, x_4) = 0\}$. Now observe, also by construction, that f is a non-constant polynomial function. Therefore, Z is closed and has zero Lebesgue measure, being the set of zeros of a non-constant polynomial function.

We use this claim to establish that $C^{(ijk\ell)}_{[abc]}$ is closed and has zero Lebesgue measure. If

$d = 3$, this is an immediate consequence of the above. So suppose that $d > 3$. Observe that whether a profile $(x_i^*)_{i=1,\dots,n,A} \in \mathbb{R}^{d(n+1)}$ is an element of $C_{[abc]}^{(ijk\ell)}$ is determined exclusively by the collection of projections $\{[x_i^*]_{abc}, [x_j^*]_{abc}, [x_k^*]_{abc}, [x_\ell^*]_{abc}\}$. Hence,

$$C_{[abc]}^{(ijk\ell)} = K \times \mathbb{R}^{d(n+1)-12}, \text{ where } K \subseteq \mathbb{R}^{3\times 4} \text{ is defined by}$$
$$K := \left\{ ([x_\nu^*]_{abc})_{\nu=i,\ell,j,k} \in \mathbb{R}^{3\times 4} : [x_i^*]_{abc}, [x_j^*]_{abc}, [x_k^*]_{abc}, [x_\ell^*]_{abc} \text{ are coplanar in } \mathbb{R}^3 \right\}.$$

Observe that K is equivalent (modulo relabeling of indices) to Z in (22), and therefore is closed and has zero Lebesgue measure in $\mathbb{R}^{12}$. Hence, $C_{[abc]}^{(ijk\ell)}$ is also closed and has zero Lebesgue measure in $\mathbb{R}^{d(n+1)}$.

The above establishes that C is closed and has zero Lebesgue measure. Therefore, its complementary set $NC := \mathbb{R}^{d(n+1)} \backslash C$, the set of ideal point profiles satisfying Non-Coplanarity, is open-dense and has full Lebesgue measure (as any open full-measure set is dense).

B.4 Failure of Manipulability in Two-Dimensional Spatial Politics

Using a three-voter example, we illustrate the assertion from Section 4.1 that, when there are $d = 2$ policy dimensions, manipulability fails whenever the agenda setter's ideal point lies outside the convex hull of voter ideal points; this analysis straightforwardly extends to a general (odd) number of voters, provided that no 3 of their ideal points are collinear (which is generically satisfied).

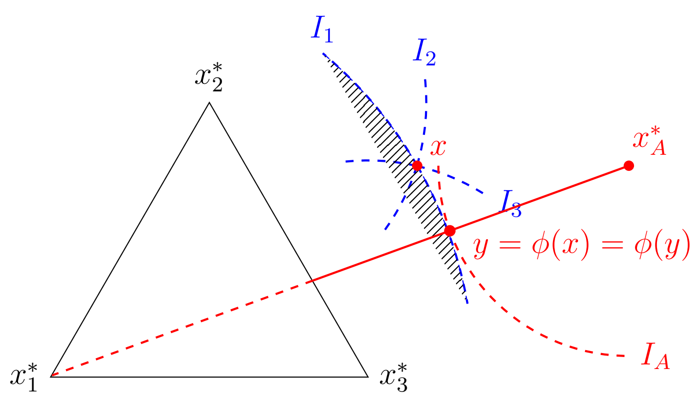

FIGURE 5. A failure of manipulability in the two-dimensional spatial model.

Consider the situation depicted in Figure 5 where $(x_i^*)_{i=1,2,3,A}$ depicts the profile of ideal points. We first observe that all policies on the line segment between x_1^* and x_A^* and outside the interior of the convex hull of voter ideal points—this is the solid red line—are unimprovable. To see why, note that for any such policy—such as policies x and y in the figure—a majority of voters favor another policy, say z, to y only if voter 1 also favors z to y.[9] Because voter 1's indifference curve passing through y is tangent to the agenda setter's indifference curve passing through y, there is no policy that the agenda setter and voter 1 both prefer to y. Thus, y is unimprovable.

Necessarily, this eliminates any prospect for a dictatorial power result starting from *any* default option: beginning with a default option of y implies that it is the unique Non-Capricious equilibrium outcome, since the agenda setter's unique favorite improvement at this policy is y itself.[10] Interestingly, it also prevents the agenda setter from fully exploiting real-time agenda control even from some improvable default options (off this line segment). For example, suppose x is the initial default option. The agenda setter's unique favorite improvement from x is the unimprovable policy y, which implies that regardless of the number of rounds, the unique Non-Capricious equilibrium outcome is y. This logic does not merely apply to the policies x and y, but is more general: there exists an open set of policies such that if the initial default option belongs to this open set, the unique Non-Capricious equilibrium outcome is bounded away from the agenda setter's ideal point. Thus, even though the set of unimprovable policies is measure-0 in $\mathbb{R}^2$, strategic forces may compel negotiations to reach that set, contravening a dictatorial power result.

B.5 Proof of Theorem 6

Step 1: Equilibrium Outcomes for General Voting Rules. We first introduce notation that extends that from Appendix B.2.1 to general voting rules. Given any voting rule $\mathcal{D}$ and $x \in X$, denote the *weak* $\mathcal{D}$-acceptance set $M^{\text{w}}_{\mathcal{D}}(x) := \{y \in X : y \succsim_i x \; \forall i \in D, \; \exists D \in \mathcal{D}\}$, the *strict* $\mathcal{D}$-acceptance set $M^{\text{s}}_{\mathcal{D}}(x) := \{y \in X : y \succ_i x \; \forall i \in D, \; \exists D \in \mathcal{D}\}$, and the *almost-strict* $\mathcal{D}$-acceptance set $M^{\text{as}}_{\mathcal{D}}(x) := \text{cl}\,[M^{\text{s}}_{\mathcal{D}}(x)] \cup \{x\}$. Define the agenda setter's *favorite almost-strict $\mathcal{D}$-improvement* value function $V^{\text{as}}_A(\cdot \mid \mathcal{D}) : X \to \mathbb{R}$ by $V^{\text{as}}_A(x \mid \mathcal{D}) := \max_{y \in M^{\text{as}}_{\mathcal{D}}(x)} u_A(y)$, and her *one-round $\mathcal{D}$-improvement* correspondence $\Phi^{\text{or}}_{\mathcal{D}} : X \rightrightarrows X$ by

$$\Phi^{\text{or}}_{\mathcal{D}}(x) := \{y \in X : y \in M^{\text{w}}_{\mathcal{D}}(x) \text{ and } u_A(y) \geq V^{\text{as}}_A(x \mid \mathcal{D})\}\,. \tag{23}$$

We denote the set of Non-Capricious equilibria in the T-round game with initial default x^0 and voting rule $\mathcal{D}$ by $\Sigma^{\text{NC}}_{\mathcal{D}}(x^0, T)$. Given any $\sigma \in \Sigma^{\text{NC}}_{\mathcal{D}}(x^0, T)$, we denote the equilibrium outcome by $g^{\sigma}_T(x^0 \mid \mathcal{D})$.

[9]To put it differently, voters 2 and 3 cannot both favor z to y without voter 1 also doing so.

[10]Formally, the agenda setter's one-round improvement correspondence Φ^{or} (as defined in Equation (12)) satisfies $\Phi^{\text{or}}(y) = \{y\}$ and the above assertion follows from Lemma 5(b). With a slight abuse of notation, in Figure 5 we let $\phi(\cdot)$ denote the unique element of $\Phi^{\text{or}}(\cdot)$ at points where this correspondence is singleton-valued, and refer to this policy as the agenda setter's unique favorite improvement.

The following generalizes Lemma 5(b) in Appendix B.2.1 to arbitrary voting rules:

Lemma 9. *For any collective choice problem $\mathcal{C}$ and voting rule $\mathcal{D}$, the following holds:*

> *For any $x^0 \in X$, $T \in \mathbb{N}$, and Non-Capricious equilibrium $\sigma \in \Sigma_{\mathcal{D}}^{NC}(x^0, T)$, there exists a collection $\{\hat{\phi}_t(\cdot)\}_{\tau=1}^{T}$ of selections $\hat{\phi}_t(\cdot) \in \Phi_{\mathcal{D}}^{or}(\cdot)$ such that the equilibrium outcome is given by*
>
> $$g_T^{\sigma}(x^0 \mid \mathcal{D}) = \left[\hat{\phi}_1 \circ \hat{\phi}_2 \circ \cdots \circ \hat{\phi}_T\right](x^0)$$
>
> *and, for every $x \in X$, we have $\hat{\phi}_t(x) \sim_A \hat{\phi}_T(x)$ for all $1 \leq t \leq T$.*

The proof of Lemma 9 is identical to that of Lemma 5(b) modulo the notational adaptation described above, and hence omitted.

Step 2: Properties of Distribution Problems. We now characterize the agenda setter's favorite policies and one-round $\mathcal{D}$-improvement operator in Distribution Problems. Throughout our analysis in this Step, we restrict attention to Distribution Problems, assume that Thin Individual Indifference holds, and consider a veto-proof voting rule $\mathcal{D}$. We denote the set of weakly Pareto efficient policies by $P := \{x \in X : \nexists y \text{ such that } \forall i \in N \cup \{A\},\, y \succ_i x\}$. We let $\underline{u}_i := \min_{x \in X} u_i(x)$ denote player i's minimal utility. For any policy x, we define its *support* by $\text{supp}(x) := \{i \in N : u_i(x) > \underline{u}_i\}$, viz., the set of voters for whom x is not a least-preferred policy. The following claim demonstrates that the agenda setter's favorite policies are precisely those that are weakly Pareto efficient and leave all voters with minimal utility:

Claim 8. $X_A^* = \{x \in P : \mathit{supp}(x) = \emptyset\}$.

Proof. For any $x \notin X_A^*$, Scarcity implies that $x \notin P$ or $\text{supp}(x) \neq \emptyset$. By contraposition, it follows that $\{x \in P : \text{supp}(x) = \emptyset\} \subseteq X_A^*$. For the opposite inclusion, consider a policy $y \notin \{x \in P : \text{supp}(x) = \emptyset\}$; we establish that $y \notin X_A^*$. If $y \notin P$, then by definition there exists a strongly Pareto dominating $z \in X$, and therefore, $y \notin X_A^*$. If $\text{supp}(y) \neq \emptyset$, then by definition there exists some voter $i \in N$ such that $u_i(y) > \underline{u}_i$. Transferability then implies that there exists some $z \in X$ such that $z \succ_j x$ for all players $j \neq i$, including $j = A$; hence, $y \notin X_A^*$. By contraposition, it follows that $X_A^* \subseteq \{x \in P : \text{supp}(x) = \emptyset\}$. □

The next claim characterizes $\Phi_{\mathcal{D}}^{\text{or}}$. Given any policies $x, y \in X$, we let $L(y \mid x) := \{i \in N : y \prec_i x\}$ denote the set of voters who are *losers* if the implemented policy changes from x to y. We say that voter $i \in N$ is *minimized* at $x \in X$ if $i \notin \text{supp}(x)$.

Claim 9. *For every $x \in X$ and $y \in \Phi_{\mathcal{D}}^{or}(x)$, the following hold:*

(a) *y is weakly Pareto efficient: $y \in P$.*

(b) *Losers are minimized: $L(y \mid x) = supp(x) \backslash supp(y)$.*

(c) *Minimized voters remain minimized: $supp(y) \subseteq supp(x)$.*

(d) *Minimal winning coalition: $\nexists D \in \mathcal{D}$ and $i \in supp(y)$ such that $y \succcurlyeq_j x$ $\forall j \in D$ and $D \backslash \{i\} \in \mathcal{D}$.*

Proof. Let $x \in X$ and $y \in \Phi^{\text{or}}_{\mathcal{D}}(x)$ be given. We prove each point by contradiction. For parts (a)-(c), take as given a winning coalition $D' \in \mathcal{D}$ such that $y \succcurlyeq_i x$ for all $i \in D'$ (existence of which is guaranteed because $\Phi^{\text{or}}_{\mathcal{D}}(x) \subseteq M^{\text{w}}_{\mathcal{D}}(x)$).

For (a), suppose $y \notin P$: then there exists a policy z such that $z \succ_i y$ for every player i. Hence, $z \in M^{\text{s}}_{\mathcal{D}}(x)$ (as $z \succ_i y \succcurlyeq_i x$ for all $i \in D'$) and $z \succ_A y$, contradicting that $y \in \Phi^{\text{or}}_{\mathcal{D}}(x)$.

For (b), it suffices to establish the inclusion $L(y \mid x) \subseteq \text{supp}(x) \backslash \text{supp}(y)$, as the opposite inclusion is tautological. Observe that by definition, $L(y \mid x) \subseteq \text{supp}(x)$. Suppose towards a contradiction that there exists some voter $i \in L(y \mid x) \cap \text{supp}(y)$. It then follows, by definition of $\text{supp}(y)$, that $u_i(y) > \underline{u}_i$. Transferability implies that there exists a policy z such that $z \succ_j y$ for all players $j \neq i$, including $j = A$ and all $j \in D'$. Hence, $z \in M^{\text{s}}_{\mathcal{D}}(x)$ and $z \succ_A y$, contradicting that $y \in \Phi^{\text{or}}_{\mathcal{D}}(x)$.

For (c), suppose there exists a voter $i \in \text{supp}(y) \backslash \text{supp}(x)$. As $u_i(y) > \underline{u}_i$, Transferability implies that there exists some $z \in X$ such that $z \succ_j y$ for all players $j \neq i$, including $j = A$ and all $j \in D' \backslash \{i\}$. As $u_i(z) \geq u_i(x) = \underline{u}_i$ by definition, there are two cases. First, if $z \succ_i x$, then $z \in M^{\text{s}}_{\mathcal{D}}(x)$ and $z \succ_A y$, contradicting that $y \in \Phi^{\text{or}}_{\mathcal{D}}(x)$. Second, if $z \sim_i x$, then since $z \neq x$ (as $z \succ_A y \succcurlyeq_A x$), we have $z \in I_i(x) \backslash \{x\}$ and $d(z, x) > 0$. For any $\epsilon \in (0, d(z, x))$, we have $x \notin B_\epsilon(z)$ and hence $B_\epsilon(z) \backslash [I_i(x) \backslash \{x\}] = B_\epsilon(z) \backslash I_i(x)$. Thin Individual Indifference then implies that, for every such $\epsilon > 0$, there exists some $z' \in B_\epsilon(z) \backslash I_i(x)$. As $I_i(x) = \{x' \in X : u_i(x') = \underline{u}_i\}$, any such z' satisfies $z' \succ_i x$. Moreover, by continuity of players' preferences, there exists a small enough $\epsilon > 0$ and corresponding $z' \in B_\epsilon(z) \backslash I_i(x)$ such that $z' \succ_j y$ for all $j \neq i$. Then $z' \in M^{\text{s}}_{\mathcal{D}}(x)$ (as $z' \succ_j y \succcurlyeq_j x$ for all $j \in D' \backslash \{i\}$ and $z' \succ_i x$) and $z' \succ_A y$, contradicting that $y \in \Phi^{\text{or}}_{\mathcal{D}}(x)$.

For (d), suppose that such $D \in \mathcal{D}$ and $i \in \text{supp}(y)$ exist. Transferability implies that there exists some $z \in X$ such that $z \succ_j y$ for all $j \neq i$, including all $j \in D \backslash \{i\}$ and $j = A$. As $D \backslash \{i\} \in \mathcal{D}$, this implies that $z \in M^{\text{s}}(x)$ (as $z \succ_j y \succcurlyeq_j x$ for all $j \in D \backslash \{i\}$) and $z \succ_A y$, contradicting that $y \in \Phi^{\text{or}}_{\mathcal{D}}(x)$. □

Step 3: Main Argument for Theorem 6. By virtue of Lemma 9, the following lemma implies Theorem 6.

Lemma 10. *Suppose $\mathcal{C}$ is a Distribution Problem satisfying Thin Individual Indifference.*

(a) *If* $\mathcal{D}$ *is a quota rule with quota* $q < n$*, then for any collection* $\{\hat{\phi}_t\}_{t=1}^T$ *of selections* $\hat{\phi}_t(\cdot) \in \Phi_{\mathcal{D}}^{or}(\cdot)$*, the following holds:*

$$\textit{If } T \geq \left\lceil \frac{n}{n-q} \right\rceil, \textit{ then } \left[\hat{\phi}_1 \circ \cdots \circ \hat{\phi}_T\right](x) \in X_A^* \textit{ for all } x \in X.$$

(b) *If* $\mathcal{D}$ *is a veto-proof voting rule, then for any collection* $\{\hat{\phi}_t\}_{t=1}^T$ *of selections* $\hat{\phi}_t(\cdot) \in \Phi_{\mathcal{D}}^{or}(\cdot)$*, the following holds:*

$$\textit{If } T \geq n, \textit{ then } \left[\hat{\phi}_1 \circ \cdots \circ \hat{\phi}_T\right](x) \in X_A^* \textit{ for all } x \in X.$$

Proof. We begin by establishing part (a). Let the quota rule $\mathcal{D}$ with quota $q < n$, number of rounds $T \geq \lceil n/(n-q) \rceil$, and $\Phi_{\mathcal{D}}^{\text{or}}$-selections $\{\hat{\phi}\}_{t=1}^T$ be given. Let $x \in X$ be given and define $z_t := \left[\hat{\phi}_t \circ \cdots \circ \hat{\phi}_T\right](x)$ for all $t \in \{1, \ldots, T\}$, with $z_{T+1} := x$. We must show that $z_1 \in X_A^*$. Claim 9(a) implies that $z_t \in P$ for all t. Thus, by Claim 8, it suffices to show that $\text{supp}(z_1) = \emptyset$.

To that end, we claim that for every $t \in \{1, \ldots, T\}$,

$$\text{supp}(z_t) \subseteq \text{supp}(z_{t+1}) \quad \text{and} \quad |\text{supp}(z_{t+1}) \backslash \text{supp}(z_t)| = \min\{n-q, |\text{supp}(z_{t+1})|\}. \tag{24}$$

Observe that (24) implies that for every $t \in \{1, \ldots, T\}$,

$$\begin{aligned} |\text{supp}(z_t)| &= |\text{supp}(z_{t+1})| - \min\{n-q, |\text{supp}(z_{t+1})|\} \\ &= \max\{|\text{supp}(z_{t+1})| - (n-q), 0\} \\ &= \max\{|\text{supp}(x)| - (T+1-t)(n-q), 0\}, \end{aligned}$$

where the first and second lines are identities and the third line follows from iteratively applying the preceding lines. This implies that $|\text{supp}(z_1)| = 0$ if and only if $T \geq |\text{supp}(x)|/(n-q)$. As $n \geq |\text{supp}(x)|$ and $T \geq \lceil n/(n-q) \rceil$ by assumption, it follows that $\text{supp}(z_1) = \emptyset$.

Therefore, it suffices to prove (24). We do so by appealing to Claim 9(b)-(d), noting that $z_t = \hat{\phi}_t(z_{t+1}) \in \Phi_{\mathcal{D}}^{\text{or}}(z_{t+1})$ for all t by construction. First, Claim 9(c) directly implies that $\text{supp}(z_t) \subseteq \text{supp}(z_{t+1})$. Next, Claim 9(b) implies that $L(z_t \mid z_{t+1}) = \text{supp}(z_{t+1}) \backslash \text{supp}(z_t)$. If $|\text{supp}(z_{t+1})| = 0$, this proves the claim. So, assume that $|\text{supp}(z_{t+1})| > 0$. We assert that $|L(z_t \mid z_{t+1})| = \min\{n-q, |\text{supp}(z_{t+1})|\}$. That $|L(z_t \mid z_{t+1})| \leq |\text{supp}(z_{t+1})|$ follows from $L(z_t \mid z_{t+1}) = \text{supp}(z_{t+1}) \backslash \text{supp}(z_t)$. That $|L(z_t \mid z_{t+1})| \leq n-q$ follows from: (i) $D \in \mathcal{D}$ if and only if $|D| \geq q$ and (ii) $L(z_t \mid z_{t+1}) \subseteq N \backslash D$ for any $D \in \mathcal{D}$ such that $z_t \succcurlyeq_i z_{t+1}$ for all $i \in D$. Hence, $|L(z_t \mid z_{t+1})| \leq \min\{n-q, |\text{supp}(z_{t+1})|\}$. Suppose, towards a contradiction, that $|L(z_t \mid z_{t+1})| < \min\{n-q, |\text{supp}(z_{t+1})|\}$. Because $|L(z_t \mid z_{t+1})| < |\text{supp}(z_{t+1})|$ and $L(z_t \mid z_{t+1}) \subseteq \text{supp}(z_{t+1})$, there exists some voter $i \in \text{supp}(z_{t+1}) \backslash L(z_t \mid z_{t+1}) = \text{supp}(z_t)$. Because $|L(z_t \mid z_{t+1})| < n-q$, the set of voters $D := N \backslash L(z_t \mid z_{t+1})$ satisfies $|D| > q$, implying $D \in \mathcal{D}$ and $D \backslash \{i\} \in \mathcal{D}$. Moreover, $z_t \succsim_j z_{t+1}$ for all $j \in D$ by construction. Therefore, Claim 9(d) implies the contradiction that $z_t \notin \Phi^{\text{or}}_{\mathcal{D}}(z_{t+1})$, as desired.

This concludes the proof of the claim and thus part (a). The proof of part (b) is very similar, so we provide only a sketch. For a general veto-proof voting rule $\mathcal{D}$, the key claim is that $\text{supp}(z_t) \subseteq \text{supp}(z_{t+1})$ and $|\text{supp}(z_{t+1}) \backslash \text{supp}(z_t)| \geq \min\{1, |\text{supp}(z_{t+1})|\}$, which implies that n rounds suffices by appeals to Claim 8 and Claim 9(a), coupled with calculations similar to those below Equation (24). That $\text{supp}(z_t) \subseteq \text{supp}(z_{t+1})$ again follows directly from Claim 9(c). To show that $|\text{supp}(z_{t+1}) \backslash \text{supp}(z_t)| \geq \min\{1, |\text{supp}(z_{t+1})|\}$, it suffices to consider the case in which $\text{supp}(z_{t+1}) \neq \emptyset$. Claim 9(b) implies that $L(z_t \mid z_{t+1}) = \text{supp}(z_{t+1}) \backslash \text{supp}(z_t)$. Suppose towards a contradiction that $|\text{supp}(z_{t+1}) \backslash \text{supp}(z_t)| = 0$, which implies that (i) $\text{supp}(z_t) = \text{supp}(z_{t+1}) \neq \emptyset$ and (ii) $z_t \succsim_i z_{t+1}$ for all voters $i \in N$. By (i), there exists some voter $k \in \text{supp}(z_t)$. By (ii) and that there exists some $D \in \mathcal{D}$ with $k \notin D$ (as $\mathcal{D}$ is veto-proof), Claim 9(d) implies the desired contradiction $z_t \notin \Phi^{\text{or}}_{\mathcal{D}}(z_{t+1})$, proving part (b). □

B.6 Proof of Theorem 8

For a subset of policies $Y \subseteq X$, the following definitions are standard: Y is *internally stable* if there do not exist distinct $x, y \in Y$ such that $y \succ_M x$ and $y \succ_A x$. Y is *externally stable* if, for every $x \notin Y$, there exists some $y \in Y$ such that $y \succ_M x$ and $y \succ_A x$. Y is *stable* if it is both internally and externally stable. As shown by Diermeier and Fong (2012), there exists a unique stable set in the present setting, which we denote by V. Recall that $E = \{x \in X : x = \phi(x)\}$ denotes the set of unimprovable policies. Observe that $E \subseteq V$ since excluding any unimprovable policy would contradict the external stability of V.

Recall from Section 3.2 that $M(x) := \{y \in X : y \succ_M x \text{ or } y = x\}$. Define the agenda setter's *favorite stable improvement* $\psi(\cdot; V) : X \to V$ by

$$\{\psi(x;V)\} := \underset{y \in M(x) \bigcap V}{\arg\max}\ u_A(y) \tag{25}$$

which is well-defined because V is externally stable. By definition, for every policy x, $\phi(x) \succsim_A \psi(x;V)$ and $\phi(x) = \psi(x;V)$ if and only if $\phi(x) \in V$.

We recall the following characterizations of equilibrium outcomes and payoffs:

- Lemma 1 shows that in the T-round game with $T < \infty$, an initial default x^0 leads to the unique equilibrium outcome $\phi^T(x^0)$. The agenda setter's equilibrium payoff is denoted $U_T(x^0) := u_A(\phi^T(x^0))$.

- Diermeier and Fong (2012) and Anesi and Seidmann (2014) show that in the T-round game with $T = \infty$, an initial default x^0 leads to the unique MPE outcome $\psi(x^0; V)$.[11] The agenda setter's equilibrium payoff is denoted $U_\infty(x^0) := u_A(\psi(x^0; V))$.

Observe that these characterizations, and the fact that $\phi^{t+1}(\cdot) \succsim_A \phi^t(\cdot)$ for all t, immediately yields the initial claim in Theorem 8. They also permit us to equivalently rephrase statements (a) and (b) from Theorem 8 as:

(a) *There exist $x^0 \in X$ such that $\phi^2(x^0) \succ_A \phi(x^0) \succ_A \psi(x^0; V)$.*
(b) *For all $x^0 \in X$, $\phi^2(x^0) = \phi(x^0) = \psi(x^0; V)$.*

Thus, it suffices to show that statements (a) and (b) above are mutually exclusive and exhaustive. Mutual exclusivity is obvious; the argument below establishes exhaustiveness.

Let $R := \{x \in X : \phi(x) \in E\}$ denote the set of *at-most-once-improvable* policies. These are the default policies at which the agenda setter does not benefit from having more than a single round of proposals in the finite-horizon game because she obtains $\phi(x)$ with a single round, and $\phi(x)$ is unimprovable. By contrast, if $x \notin R$, the agenda setter strictly prefers having two or more rounds in the finite-horizon game to a single round.

We show that the following identity holds:

$$R = \{x \in X : \phi(x) = \psi(x; V) \text{ and } \phi^2(x) = \psi(\phi(x); V)\}. \tag{26}$$

To see why (26) is true, suppose that x is an element of the set on the RHS. Observe that $\phi(x) = \psi(x; V)$ implies that $\phi(x) \in V$. Because V is internally stable, it then follows that $\psi(\phi(x); V) = \phi(x)$. Therefore, $\phi^2(x) = \psi(\phi(x); V) = \phi(x)$, which implies that $\phi(x) \in E$, and therefore, $x \in R$. Proceeding in the other direction, suppose $x \in R$. By definition of R, $\phi(x) \in E$, which implies that $\phi^2(x) = \phi(x)$. Because $E \subseteq V$, it also follows that $\phi(x) \in V$ and therefore, $\phi(x) = \psi(x; V)$. As V is internally stable, $\psi(\phi(x); V) = \phi(x)$. Therefore, $\phi^2(x) = \psi(\phi(x); V)$.

We use Equation (26) to show that Statements (a) and (b) reduce to the two exhaustive cases of $R \subsetneq X$ and $R = X$. We begin with the latter.

If $R = X$, then it follows from the definition of R that for every policy x, $\phi^2(x) = \phi(x)$, as $\phi(x)$ is unimprovable; it also follows from (26) that $\phi(x) = \psi(x; V)$. Therefore, Statement (b) necessarily holds.

[11]Specifically, Propositions 1 and 2 in Anesi and Seidmann (2014), specialized to the present setting with a single proposer and Generic Finite Alternatives, imply the above characterization for *some* stable set; as noted above, Lemmas 1-3 in Diermeier and Fong (2012) show that the stable set V exists and is unique in the present setting. Theorem 1 in Diermeier and Fong (2012) provides an analogous characterization of MPE outcomes in the context of Diermeier and Fong's (2011) infinite-horizon model (with discounting and no termination rule).

If $R \subsetneq X$, then there is a policy $y \notin R$. It follows from the definition of R that $\phi(y) \notin E$, and hence $\phi^2(y) \succ_A \phi(y)$. Moreover, (26) establishes that either $\phi(y) \succ_A \psi(y;V)$ or $\phi^2(y) \succ_A \psi(\phi(y);V)$. In the former case, Statement (a) is established for $x^0 = y$. In the latter case, (26) implies that $\phi(y) \notin R$, and therefore, $\phi^2(y) \notin E$. Hence, $\phi^3(y) \succ_A \phi^2(y) \succ_A \psi(\phi(y);V)$. Statement (a) is now established for $x^0 = \phi(y)$.