Representation in Collective Policymaking

Daniel Gibbs[*] Gleason Judd[†]

October 20, 2025

Abstract

We study a policy-motivated principal choosing a representative to bargain over one-dimensional policy in a collective body under majority rule. The representative's ideology affects policy outcomes directly through their own proposal and indirectly by shaping the median's expectations, determining which proposals pass. This creates a tradeoff: a representative close to the principal's ideal point makes favorable proposals but permits extremist proposals to pass, while a more centrist representative constrains extremism by improving the median's continuation value. We show that a wide range of principals strictly prefer representatives biased toward center rather than themselves. The optimal representative is unique for almost all principals, strictly increasing in the principal's ideal point, and never biased away from center. We also study competitive representation with multiple principals choosing representatives.

Keywords: representation, collective policymaking, bargaining, spatial policy, moderation

JEL Codes: C78, D70, D72

---

[*] Corresponding Author: Virginia Tech (`gibbsd@vt.edu`).

[†] Princeton University (`gleason.judd@princeton.edu`).

1 Introduction

Representation in collective policymaking is common but complex. Many policy decisions are made in collective bodies by representatives, such as legislators representing their constituents, committee members representing their party leaders, or judges who politicians appoint to multi-member courts. Since those policies can depend on the overall composition of representatives, as well as their institutional rights and roles (Romer and Rosenthal 1978; Baron and Ferejohn 1989; Krehbiel 1998), the impact of individual representatives can be subtle (Miller and Stokes 1963; Eulau and Karps 1977).[1]

We study the impact and appeal of individual representatives in collective bodies. How do individual representatives impact collective policymaking? Which kinds of representatives are optimal? How does one's optimal representative depend on the characteristics of other members of the collective body or their institutional rights?

We aim to sharpen theoretical understanding by accounting for two key complications of collective policymaking. First, it is interdependent, as individual representatives may impact each other in a variety of ways (Harstad 2010; Gailmard and Hammond 2011). Second, it is uncertain, as forecasts about things like the duration, proposals, or outcomes are typically noisy (Fowler 2006). These two complications can have common sources—e.g., voting rules, procedural rights, agenda congestion, ideological heterogeneity, or polarization—and also impact each other. Although these features and their connection have been incorporated into models of collective policymaking (e.g., Baron and Ferejohn 1987, 1989; Baron 1996; Banks and Duggan 2000, 2006), their consequences for representation are undeveloped.

We analyze a policy-motivated principal choosing the ideal point of their representative, who will bargain with other politicians over one-dimensional policy under simple majority rule. We highlight how different representatives not only behave differently but also induce some of the other politicians to behave differently. The representative's expected behavior impacts which policies can pass and, in turn, affects proposals by extreme politicians. We show that a broad range of principals want to bias their representative inwards, to improve expectations of (de facto) veto players and further constrain extremists. Thus, we find a widespread aversion against more extreme representatives.

Our collective policymaking setting consists of sequential bargaining over an infinite horizon à la Banks and Duggan (2000). In each period until agreement, a politician is recognized to propose a policy from a one-dimensional policy space and then a simple majoritarian vote determines whether their proposal passes or bargaining continues. All players are policy-motivated, with preferences over policy represented by quadratic loss in a *bad status quo* setting where any agreement is preferable to the status quo. The key heterogeneity between politicians is in their ideal points, but we also allow them to have different proposal rights (i.e., recognition probabilities). We fix those proposal rights, however, so that the principal can only choose the representative's ideal point. Accordingly, we isolate the impacts of individual ideological differences between representatives.

Once a representative is in place, equilibrium policymaking induces a unique lottery over policy (Cho and Duggan 2003; Cardona and Ponsati 2011). Whoever proposes first will pass their favorite policy in the set that a majority would pass (Banks and Duggan 2000). Furthermore, that set always coincides with the median politician's *acceptance set* (Duggan 2014), which is an interval of policies around her ideal point.

---

[1] Miller and Stokes (1963) highlights that "[t]he legislator acts in a complex institutional setting in which he is subject to a wide variety of influences" (pg. 51) and Eulau and Karps (1977) echoes that "[. . .] representatives are influenced in their conduct by many forces or pressures or linkages [. . .]" (pg. 235). More broadly, Pitkin (1967) states that "representation is not any single action by any one participant, but the overall structure and functioning of the system, the patterns emerging from the multiple activities of many people" (pg. 221).

Crucially, it is determined by her expectations about further policymaking—since policymaking can continue after rejected proposals—and thus depends on the profile of politician ideal points and their proposal rights.

The representative's ideology can indirectly affect what some of the other politicians propose in equilibrium. It does so by changing the acceptance set, through shifting either the median's (i) location *or* (ii) policy expectations. The location channel is familiar (Gailmard and Hammond 2011; Klumpp 2010), but the policy expectations channel is less understood. In our setting, the latter is especially pervasive. Regardless of whether the representative would be the median, his mere presence affects the median's willingness to reject proposals due to the prospect that he might subsequently propose. Essentially, the representative's ideology can have *anticipation effects* (Friedrich 1937; Simon 1953). We parse these two channels and characterize how the acceptance set varies with the representative's ideal point. Over representatives who would be the median, it shifts monotonically with the representative's ideal point but its radius can change in different ways, depending on the distribution of proposal rights. Over representatives who would not be the median but are sufficiently moderate that the median accepts the representative's ideal policy if proposed, the acceptance set contracts as the representative shifts towards the median since the median's continuation value from rejecting improves. The representative's ideology has no marginal effect on the distribution over policy outcomes only if the representative is sufficiently extreme such that the representative's ideal policy is rejected by the median if proposed.

Due to the representative's indirect impact on other politicians, the principal faces a classic tradeoff: although a biased representative may propose less favorable policy, they may also induce others to propose more favorable policies (Schelling 1956). Appointing a representative near her own ideal point ensures favorable proposals if that representative is recognized, but leaves the (decisive) median politician with weak prospects from continued bargaining, inducing them to accept a wide range of proposals. In contrast, appointing a more centrist representative would shift the representative's proposal away from the principal but substantially improve the median's continuation value, inducing them to accept a narrower range of proposals. This tightens the constraint on extremists. We show that this indirect benefit—constraining extremist proposals through the median's expectations—often dominates the direct cost of the representative's less favorable proposal.

To illustrate, consider a stylized five-politician setting: two politicians at the far left and far right (positions 0 and 1), two at intermediate positions (0.4 and 0.6), and a principal at 0.3 who is choosing a representative. With sufficient patience, the equilibrium exhibits the following pattern. If the principal appoints a representative at 0.3, the acceptance set is approximately $[0.2, 0.7]$, with the far-left and far-right politicians proposing at these boundaries. If instead she appoints a centrist at 0.5 who becomes median, the acceptance set tightens to approximately $[0.3, 0.6]$. The representative's proposal shifts from 0.3 to 0.5, but proposals from both the far-left and far-right moderate by 0.1 each. Thus, the principal faces a tradeoff: a less favorable proposal by the representative, but a tighter constraint on extreme proposals.

We highlight how this tradeoff arises more generally and may be more widespread than previously appreciated. It arises from pervasive anticipation effects rather than requiring the representative to be a veto player or the status quo to be strategically relevant.

We characterize general properties of optimal representatives under minimal assumptions on politician ideal points and recognition probabilities and show that the principal is always more inclined towards moderation than extremism. The principal never strictly prefers someone more extreme and instead always wants someone who is the median or biased in that direction. More precisely, very extreme principals are indifferent among all extremist representatives from their side, so they can self-represent but have no strict preference to do so. Next, if the principal is in either of two intermediate intervals flanking a central interval, then their optimal representative is strictly more centrist than themselves but not the median—specifically, choosing from the moderate region on their side. For principals in a centrally located interval, their optimal representative is the median (a centrist). For any principal who chooses a moderate representative, the downside of biasing their representative's proposal is outweighed by the upside of inducing extremists to further moderate their proposals. Additionally, optimal representatives are ordered: more right-leaning principals choose more right-leaning representatives. Thus, the principals who want a centrist representative form an interval, surrounded by two intervals of principals who each want a moderate representative more centrist than themselves but not a centrist.

We analyze several variants of our general model to refine our results and illustrate potential applications. We first consider a setting with polarized legislatures, where the other members of the collective body consist of two moderates (e.g., center-left and center-right factions) and two extremists from opposite ends of the policy spectrum. We show that all non-extremist principals bias toward a unique central location characterized by the balance of extremist proposal rights. That location, the *locus of attraction*, identifies the unique principal who strictly prefers an unbiased representative.[2] Additionally, there are always *dead zones*—open intervals of representatives who are not optimal for any principal—around at least one boundary between moderate and centrist regions. Furthermore, optimal representatives vary with the balance of extremist proposal rights: principals bias further away from the empowered side to better constrain those extremists.

We then study *competitive representation*, extending our baseline to have multiple principals simultaneously choosing their representatives' ideologies. We begin with a two-principal case that reveals key strategic forces: substitution effects (where one principal's moderation reduces the other's incentive to moderate) and complementarity effects (where moderation by one principal increases the acceptance set's responsiveness to the other's choice). Which effect dominates depends on the balance of extremist proposal power. We show that strategic moderation is robust, with both principals choosing moderate representatives, though they may moderate more or less than in the single-principal benchmark depending on extremist power. The qualitative pattern extends to n principals: in any equilibrium, all principals bias strictly toward the median principal's representative, which serves as the unique locus of attraction. This differs from both the single-principal model (where the locus depends on extremist power balance) and the two-principal case (where both moderate toward a fixed median).

1.1 Contributions to the Literature

Our results provide insight into representation across various collective policymaking contexts. Our model of collective policymaking is a *minimal legislative process* (Baron 1994) with several interpretations.[3] It provides a lens for studying representation in separation-of-powers systems[4] (Epstein and O'Halloran 2001; Volden 2002) or congressional committees.[5] Broadly, we emphasize the role of *ideological* factors for representation,

---

[2] An unbiased representative is weakly optimal for very extreme principals.

[3] For discussion of interpretations and applications of our bargaining environment, see, e.g., Baron and Ferejohn (1989); Baron (1991); McCarty (2000); Kalandrakis (2006); Eraslan and Evdokimov (2019).

[4] In this vein, we add to Gailmard and Hammond (2011) and Epstein and O'Halloran (2001), who suggest "that theories of legislative organization should be brought out of the legislature and seen as part of our larger constitutional system of policy-making" (pg. 391).

[5] For an overview of scholarship in committee composition, see Evans (2011). Theoretical work on committees has studied, e.g., their *representativeness* (Krehbiel 1990; Hall and Grofman 1990; Cox and McCubbins 2007), who serves on them (Rohde and Shepsle 1973), and the role of intercameral considerations (Diermeier and Myerson 1999; Gailmard and Hammond 2011).

complementing related work emphasizing *distributive* or *informational* factors.[6]

We shed new light on how biased representatives can provide a useful form of commitment (Schelling 1956; Sobel 1981) to improve *other* politicians' behavior enough to outweigh their *own* less-favorable behavior (e.g., Harstad 2010; Christiansen 2013; Loeper 2017).[7] One prominent mechanism is that a status-quo-biased representative with veto power constrains extreme proposals (Gailmard and Hammond 2011; Klumpp 2010). Our setting includes that mechanism but also features a different mechanism—the representative's effect on expectations about policymaking—that is present in well-known models of collective policymaking (e.g., Banks and Duggan 2000). Since both mechanisms may be present in various settings (e.g., Banks and Duggan 2006), our results complement earlier work by showing how this strategic tension does not require the representative to be a veto player nor the status quo to be strategically relevant.

We highlight a new logic for how moderate representatives can appeal by reducing extremism, which complements related mechanisms in the literature. First, when allocating proposal rights, risk-averse politicians share an aversion to egalitarianism and would rather shift proposal rights towards moderate members—to make extreme proposals *less likely* (Diermeier et al. 2020).[8] In contrast, we fix (possibly unequal) proposal rights and show a widespread preference for relatively centrist representatives, to make extreme proposals *less extreme*.[9] Second, during bargaining where accepted policy becomes the new status quo, proposers may opt for centrist policies that increase the median's reservation value in *future* periods, constraining opposition *in the future* (Baron 1996; Buisseret and Bernhardt 2017; Zápal 2020). In contrast, in our setting a more centrist representative increases the median's reservation value *today*, thereby constraining what extremists can pass *today*. Moreover, in our analysis, moderate principals want to constrain extremists on both sides, not just their opponents. Third, interest groups seeking access may prefer to target more extreme representatives to moderate their proposals, improving centrist expectations and constraining what extremists can pass (Judd 2023). Our results highlight that when representatives are chosen, similar incentives encourage those groups to support selecting more moderate candidates.

Our moderation results contrast with the strategic delegation literature in Nash bargaining and conflict games, where principals often benefit from appointing more extreme agents (e.g., Schelling 1956; Sobel 1981; Segendorff 1998). In those settings, delegating to a tougher negotiator serves as a commitment device in bilateral or competitive interactions where extremism provides commitment value or competitive advantage. Our setting differs fundamentally: multilateral collective decision-making with sequential bargaining under majority rule, where a representative's ideology affects not only their own proposal but also shapes other politicians' proposals through anticipation effects on the median's voting calculus. This institutional structure—particularly the presence of multiple decision-makers under majority rule—induces continuation values that depend on the full distribution of ideal points, generating widespread incentives for moderation rather than extremism.

The moderation incentives we uncover complement extremism incentives that arise in other settings with

---

[6] Echoing Fenno (1974), Epstein and O'Halloran (2001) claim that "each of the distributive, informational, and partisan theories predicts outcomes accurately in its own relevant domain; [...] so alternative explanations should be seen as complements rather than substitutes" (pg. 391).

[7] Also see, e.g., Persson and Tabellini (1992); Besley and Coate (2003). For a general overview of strategic pre-commitment in bargaining, see Miettinen (2022). For related strategic delegation incentives outside the political context, see Dixit (1980); Vickers (1985); Bulow et al. (1985) and Fershtman et al. (1991).

[8] For other related work on *endogenous procedures*, see Diermeier and Vlaicu (2011); Diermeier et al. (2015, 2016). In a dynamic setting with endogenous status quo, Duggan and Kalandrakis (2012) endogenize proposal rights but only study equilibrium existence. In a setting with distributive policy, Eguia and Shepsle (2015) endogenize the set of politicians and their proposal rights.

[9] Two other differences, motivated by our representative/delegate application, are that in our analysis (i) the location of the median policymaker can shift and (ii) we vary the principal's ideal point.

collective policymaking. Under supermajority rule with take-it-or-leave-it proposals, voters may prefer a more extreme representative who becomes a veto pivot (Kang 2017). Where policy is a weighted average of politician ideal points, extreme representatives can counterbalance extreme opponents (Alesina and Rosenthal 1996; Kedar 2005, 2009). When principals care about how their representative votes on an *exogenous* legislative agenda, preferences can favor extremism (Patty and Penn 2019).[10]

2 Model

Players. There is a principal, P; a continuum of potential representatives; and a set of k (even) auxiliary politicians, K.

Timing. The model has two stages. First, in the *appointment* stage, P selects a representative, denoted d, to bargain on her behalf. Second, in the *bargaining* stage, the representative d interacts with the other politicians in K to collectively set a one-dimensional policy. Each bargaining period $t \in \{1, 2, ...\}$, a politician $i \in N = K \cup \{d\}$ is drawn from the recognition distribution ρ, where $\rho_i \in (0, 1)$ for all i and $\sum_{i \in N} \rho_i = 1$, and then i proposes a policy $x^t \in X = [0, 1]$. Next, all politicians vote to accept or reject x^t. The proposal is approved if and only if a simple majority of individuals approve. If x^t is approved, then it is implemented and the game ends. Otherwise, the proposal is rejected and the game moves to $t + 1$. Bargaining continues indefinitely until a proposal is accepted.

Preferences. All players are purely policy-motivated and each player has a unique ideal point $y_i \in X$. We denote the principal's ideal point as y_p and the ideal point of her chosen representative as y_d. The ideal points of the k politicians in K are ordered such that $y_1 \leq y_2 \leq ... \leq y_k$, and we denote $\ell = \frac{k}{2}$ and $r = \frac{k}{2} + 1$. The median politician in N depends on y_d and is denoted

$$m = \begin{cases} \ell & \text{if} \quad y_d < y_\ell \\ d & \text{if} \quad y_d \in [y_\ell, y_r] \\ r & \text{if} \quad y_d > y_r. \end{cases} \tag{1}$$

Once a policy x is enacted, player i will receive policy utility $u(x, y_i) = 1 - (x - y_i)^2 \geq 0$ each period thereafter. Before then, every player receives zero utility in each period until agreement.

Cumulative payoffs are sums of per-period utilities, discounted by the common factor $\delta \in (0, 1)$. We normalize per-period utility by the factor $1 - \delta$. Thus, if x is accepted in period t, then politician i's payoff is $\delta^{t-1} u(x, y_i)$.

Information. All features of the game are common knowledge.

Strategies & Equilibrium concept. In the appointment stage, a pure strategy for the principal prescribes a choice of d's ideal point, $y_d \in X$. We focus on a standard class of bargaining strategies (Banks and Duggan 2000; Cardona and Ponsati 2011) that are relatively simple and focal (Baron and Kalai 1993; Baron 1994), with politicians always voting as if pivotal (Duggan and Fey 2006). That is, when deciding whether to accept a proposal, each politician votes for the proposal if and only if they prefer it to their expected continuation payoff, treating their own vote as decisive. In the bargaining stage, a pure *stationary strategy* for each individual $i \in N$ prescribes (i) a proposal, x_i, that he makes at any t he is selected to propose; and (ii) an acceptance set, A_i, that specifies a time-independent set of proposals that he accepts or rejects. A

---

[10] Extremism can also emerge if the principal does not know their appointee's ideology but does know they will serve on a collective body that sets policy at the median ideal point (Bailey and Spitzer 2018).

stationary subgame perfect equilibrium in the bargaining subgame is a profile of stationary strategies that are mutual best responses in each subgame of the bargaining subgame. An *equilibrium* is a strategy profile in which (i) players in the bargaining subgame play stationary subgame perfect equilibrium strategies and (ii) P chooses y_d to maximize her expected payoff anticipating the distribution of policy outcomes that y_d will induce.

3 Analysis

We first characterize equilibrium behavior during the bargaining stage, after y_d is chosen. Then, we trace how y_d affects both d's behavior and that of other politicians. Next, we study the principal's preference over y_d and how her set of optimal representatives varies with her ideology. Finally, we study several special cases and extensions.

4 Equilibrium policymaking

In equilibrium, policymaking behavior is straightforward: bargaining ends immediately, with the first proposer proposing their favorite policy among those that will pass (Banks and Duggan 2000; Cardona and Ponsati 2011). That is, equilibrium play features *no delay* and the policy outcome depends only on which politician is recognized. Each politician always prefers to propose the closest majority-approved policy rather than delay. Thus, equilibrium policymaking is characterized by the *acceptance set* of policies that would pass if proposed. That set is uniquely determined by the voting calculus of the politician with the median ideal point, denoted y_m.

The median politician will pass any proposal that they prefer over their continuation value from rejecting and continuing to bargain, denoted V_m. Consequently, the acceptance set is an interval $A(y_d) = [\underline{x}(y_d), \overline{x}(y_d)]$ centered at y_m with a radius that depends on V_m. In equilibrium, V_m depends on m's expectations about the policies that other politicians would propose if recognized. Specifically, given an acceptance interval $[\underline{x}, \overline{x}]$:

$$V_m(\underline{x}, \overline{x}) = P(\underline{x})u(\underline{x}, y_m) + (1 - P(\overline{x}))u(\overline{x}, y_m) + \sum_{j \in N : y_j \in (\underline{x}, \overline{x}]} \rho_j u(y_j, y_m), \tag{2}$$

where $P(x) \equiv \sum_{i \in N : y_i \leq x} \rho_i$ denotes the cumulative proposal rights of politicians left of x. Thus, $V_m(\underline{x}, \overline{x})$ is determined by: (i) proposals at the lower bound $\underline{x}$ by politicians with $y_i \leq \underline{x}$ (weighted by their aggregate proposal rights), (ii) proposals at the upper bound $\overline{x}$ by politicians with $y_i > \overline{x}$, and (iii) proposals at ideal points $y_j \in (\underline{x}, \overline{x}]$ by interior politicians. This characterization is consistent with each politician i's equilibrium proposal strategy specifying the unique acceptable policy closest to y_i.

The equilibrium acceptance set is therefore characterized by $\underline{x}(y_d) = \min\{x \in X \mid x < y_m \text{ and } u(x, y_m) \geq \delta V_m(\underline{x}(y_d), \overline{x}(y_d))\}$, which is the leftmost proposal that m would pass, with $\overline{x}(y_d)$ defined analogously for $x > y_m$. Furthermore, since the utility function u is quadratic, the acceptance set admits a simple characterization when it is interior, with $\overline{x}(y_d) = 2y_m - \underline{x}(y_d)$ and

$$\underline{x}(y_d) = x \in (0, y_m) \text{ such that } u(x, y_m) = \delta V_m([\underline{x}(y_d), 2y_m - \underline{x}(y_d)]). \tag{3}$$

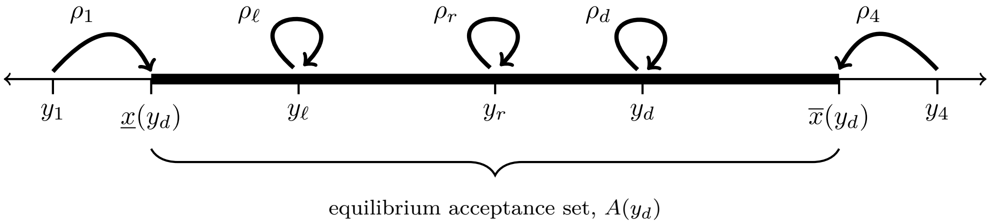
Lemma 1 formalizes key properties of the equilibrium characterization and Figure 1 illustrates.

Lemma 1. *For each $y_d \in X$, the following hold:*

1. *An equilibrium exists and it is a no-delay pure strategy equilibrium.*
2. *There is a unique equilibrium acceptance set, $A(y_d)$, of proposals that are approved if proposed and each politician i proposes the policy $x \in A(y_d)$ that minimizes $|x - y_i|$. The equilibrium acceptance set is an interval, $A(y_d) = [\underline{x}(y_d), \overline{x}(y_d)]$ and is equivalent to the set of proposals accepted by the median politician.*
3. *A $\underline{\delta}_{y_d} < 1$ exists such that $A(y_d) \subset (0, 1)$ if and only if $\delta > \underline{\delta}_{y_d}$.*

Moreover, $A(y_d)$ is continuous in y_d, δ, and ρ.

Figure 1: Illustration of equilibrium policymaking (given y_d)



Note: Figure 1 illustrates Lemma 1 for a hypothetical five-member legislature with $y_d > y_r$. The acceptance set is the bold interval, which is centered around $y_m = y_r$. Arrows point from politicians to their proposals (if recognized). Each politician proposes the closest acceptable policy.

A key implication of Lemma 1 is that every potential representative induces a well-defined policy lottery. This lottery depends on the representative through both d's own proposal (if recognized) and how d's presence affects the acceptance set. Specifically, y_d can affect: the proposals of other politicians who are constrained by $A(y_d)$ via $\underline{x}(y_d)$ and $\overline{x}(y_d)$, and the probability of constrained proposals via $P(\underline{x}(y_d))$ and $1 - P(\overline{x}(y_d))$. Remark 1 collects these observations.

Remark 1. *Given y_d, the unique equilibrium policy lottery puts probability $P(\underline{x}(y_d))$ on $\underline{x}(y_d)$; $1 - P(\overline{x}(y_d))$ on $\overline{x}(y_d)$; ρ_i on each y_i in $(\underline{x}(y_d), \overline{x}(y_d)]$; and zero otherwise.*

To simplify analysis, we focus on cases in which $A(y_d) \subset (0, 1)$ for all y_d by assuming players are sufficiently patient. The third part of Lemma 1 guarantees this is possible: for any given y_d, the acceptance set converges to $\{y_m\}$ as $\delta \to 1$.[11] Thus, $A(y_d) \subset (0, 1)$ for all y_d if $\delta > \underline{\delta} \equiv \max \underline{\delta}_{y_d}$.

Assumption 1. *$\delta > \underline{\delta}$.*

5 The Representative's Effects on Policymaking

We now characterize how the representative's ideal point affects the acceptance set and thereby the proposals of all politicians. The representative influences policy outcomes through two channels: (i) their own proposal if recognized, and (ii) the proposals of others through how y_d shapes the acceptance set $A(y_d)$. The acceptance set varies with y_d via shifts in its center (through changes in the median's identity, y_m) and its radius (through changes in the median's continuation value, V_m).

---

[11] This result follows from Banks and Duggan (2000).

The representative's proposal varies with y_d if and only if it is not constrained by the acceptance set—that is, $y_d \in \text{int}A(y_d)$. Lemma 2 establishes that there is a unique open interval $(\underline{x}_\ell, \overline{x}_r)$ such that representatives with y_d in this interval propose their ideal point if recognized, while those outside are constrained to propose the nearest boundary of the acceptance set.

Lemma 2. *There are unique* $\underline{x}_\ell \in (0, y_\ell)$ *and* $\overline{x}_r \in (y_r, 1)$ *such that* $y_d \in \text{int}A(y_d)$ *if and only if* $y_d \in (\underline{x}_\ell, \overline{x}_r)$.

The key insight is to identify the boundaries beyond which a representative becomes "too extreme" to propose their own ideal point. First, we characterize the acceptance sets at the extreme cases where $y_d = 0$ and $y_d = 1$—these define the boundaries $\underline{x}_\ell$ and $\overline{x}_r$. Second, any representative beyond these boundaries remains constrained: if shifting y_d all the way to an extreme makes them constrained, moving even further cannot make them unconstrained. Third, by continuity of the acceptance set (Lemma 1), any representative between these boundaries lies strictly inside some acceptance set.

More precisely, consider when $y_d = 0$. The representative proposes $\underline{x}(y_d)$ (since they cannot propose anything further left) and the median is $m = \ell$. Let $A_\ell = [\underline{x}_\ell, \overline{x}_\ell]$ denote the unique equilibrium acceptance set in this case, where all $i \in K$ propose $x_i^* = \text{argmin}_{x \in A_\ell} |y_i - x|$ and d proposes $\underline{x}_\ell$. Crucially, since d's proposal affects the median's continuation value only through the boundary $\underline{x}$, this same acceptance set A_ℓ must be the equilibrium for all $y_d \leq \underline{x}_\ell$—all such representatives are "bunched" at proposing the same boundary. Analogously, when $y_d = 1$, there is a unique equilibrium acceptance set $A_r = [\underline{x}_r, \overline{x}_r]$ that applies for all $y_d \geq \overline{x}_r$. Since in equilibrium d proposes $\underline{x}(y_d)$ if and only if $y_d \leq \underline{x}(y_d)$ and proposes $\overline{x}(y_d)$ if and only if $y_d \geq \overline{x}(y_d)$, it follows that $y_d \in \text{int}A(y_d)$ if and only if $y_d \in (\underline{x}_\ell, \overline{x}_r)$.

Since $[y_\ell, y_r] \subset (\underline{x}_\ell, \overline{x}_r)$, we can partition the set of potential representatives based on whether they would be the median politician and whether they propose a boundary of the acceptance set or their own ideal point. We label these three cases in Definition 1.

Definition 1. *A player i is **centrist** if $y_i \in [y_\ell, y_r]$; **extremist** if $y_i \notin (\underline{x}_\ell, \overline{x}_r)$; and **moderate** otherwise.*

These distinctions determine how y_d affects the acceptance set $A(y_d)$. The key question is whether y_d shifts the center (y_m), the radius (via V_m), or both—which hinges on whether the representative is centrist, moderate, or extremist.

Extremists have no marginal impact on the acceptance set. Any extremist is constrained by the acceptance set. The median politician remains constant on each extremist interval (either y_ℓ or y_r), and all extremist representatives on the same side propose the same boundary of $A(y_d)$. Thus both y_m and V_m are constant over each extremist interval, implying $A(y_d) = [\underline{x}_\ell, \overline{x}_\ell]$ for all left-extremists and $A(y_d) = [\underline{x}_r, \overline{x}_r]$ for all right-extremists.

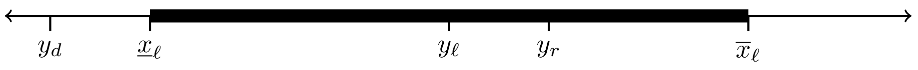
Moderates only affect the acceptance set's radius. Any moderate is not the median but, if recognized, would pass their ideal point. The median remains constant at either y_ℓ or y_r, but as y_d moves toward the median, the representative's proposal moves closer to y_m, strictly increasing V_m. A higher continuation value makes the median less willing to accept extreme proposals, so the acceptance set contracts with both bounds shifting inward at equal rates. These changes vanish as y_d approaches the centrist interval: since u is strictly concave and differentiable, the effect of d's proposal on V_m converges to zero as $y_d \to y_m$.

Centrists affect both the acceptance set's center and its radius. Any centrist representative is the median and, if recognized, would pass their ideal point, directly shifting y_m. This also changes the median's valuation of other politicians' proposals, affecting V_m and thus the radius. Although these effects can potentially oppose each other (as y_m shifts relative to other potential proposers), the direct effect through shifting y_m always dominates. The acceptance set therefore shifts in the same direction as the representative, with both boundaries strictly increasing as y_d increases over the interval of centrists.

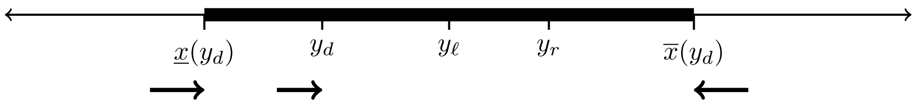
Lemma 3 formalizes these observations and establishes additional properties, with Figure 2 illustrating the key mechanisms.

Figure 2: How the representative affects policymaking

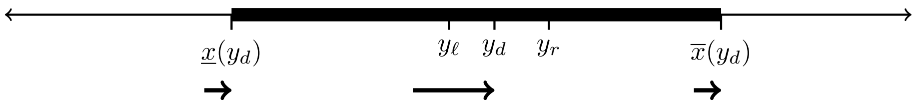
(a) all left-extremist y_d induce $A(y_d) = [\underline{x}_\ell, \overline{x}_\ell]$:



(b) as left-moderate y_d shifts inwards, $A(y_d)$ shrinks on both sides:



(c) as centrist y_d shift rightward, $A(y_d)$ also shifts rightward:



Note: Figure 2 illustrates how $A(y_d)$ changes over $y_d \leq y_r$: (a) all left-extremist y_d induce $A(y_d) = [\underline{x}_\ell, \overline{x}_\ell]$; (b) as left-moderate y_d shift inwards, $A(y_d)$ shrinks on both sides; and (c) as centrist y_d shift rightward, $A(y_d)$ also shifts rightward. Analogous properties hold for $y_d > y_r$.

Lemma 3. *The acceptance set is continuous in* y_d *and satisfies* $A(y_d) = [\underline{x}(y_d), \overline{x}(y_d)] \subset [\underline{x}_\ell, \overline{x}_r]$ *for all* y_d*. Moreover:*

1. *for all left extremist representatives* $y_d \leq \underline{x}_\ell$*, the acceptance set* $A(y_d) = [\underline{x}_\ell, \overline{x}_\ell]$ *is constant;*

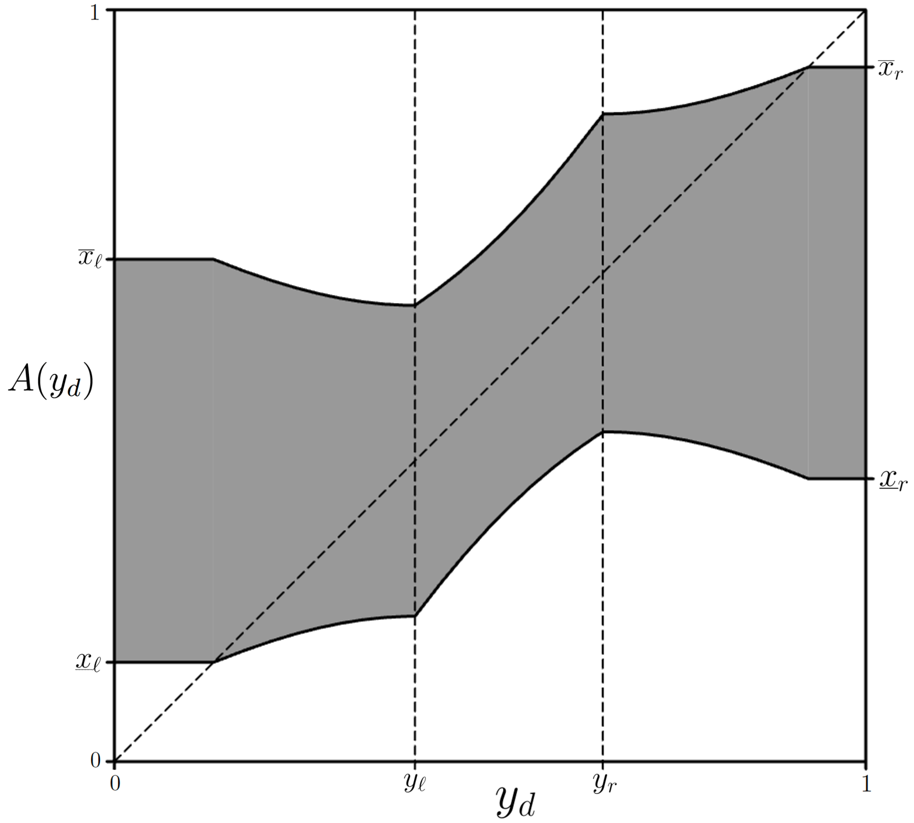
2. *for left-moderate representatives* $y_d \in (\underline{x}_\ell, y_\ell)$*, the acceptance set* $A(y_d) \subset [\underline{x}_\ell, \overline{x}_\ell]$ *contracts as* y_d *increases, with* $\underline{x}(y_d)$ *strictly increasing and* $\overline{x}(y_d)$ *strictly decreasing at equal rates that converge to zero as* $y_d \to y_\ell$*;*

3. *for centrist representatives* $y_d \in [y_\ell, y_r]$*, the acceptance set* $A(y_d) \subset [\underline{x}(y_\ell), \overline{x}(y_r)]$ *shifts rightward, with both* $\underline{x}(y_d)$ *and* $\overline{x}(y_d)$ *strictly increasing;*

4. *for right-moderate representatives* $y_d \in (y_r, \overline{x}_r)$*, the acceptance set* $A(y_d) \subset [\underline{x}_r, \overline{x}_r]$ *expands as* y_d *increases, with* $\underline{x}(y_d)$ *strictly decreasing and* $\overline{x}(y_d)$ *strictly increasing at equal rates that converge to zero as* $y_d \to y_r$*; and*

5. *for all right extremist representatives* $y_d \geq \overline{x}_r$*, the acceptance set* $A(y_d) = [\underline{x}_r, \overline{x}_r]$ *is constant.*

Figure 3: How the acceptance set varies with y_d



Note: Figure 3 illustrates how the acceptance set $A(y_d)$ (vertical axis) varies with the representative's ideal point y_d (horizontal axis) for a five-member legislature. The acceptance set (i) is constant at $[\underline{x}_\ell, \overline{x}_\ell]$ for all left-extremists ($y_d \leq \underline{x}_\ell$); (ii) contracts inward over left-moderates ($y_d \in (\underline{x}_\ell, y_\ell)$) at a decreasing rate that approaches zero as y_d nears y_ℓ; (iii) shifts rightward over centrists ($y_d \in [y_\ell, y_r]$) with both boundaries strictly increasing; (iv) expands outward over right-moderates ($y_d \in (y_r, \overline{x}_r)$) at an increasing rate starting from zero near y_r; and (v) is constant at $[\underline{x}_r, \overline{x}_r]$ for all right-extremists ($y_d \geq \overline{x}_r$). The figure uses $\delta = .98$, $(\rho_1, ..., \rho_4, \rho_d) = (.2, .15, .2, .18, .27)$, and $(y_1, ..., y_4) = (0, .4, .65, 1)$.

Additionally, $\underline{x}_\ell \leq \underline{x}_r$ *and* $\overline{x}_\ell \leq \overline{x}_r$.

Lemma 3 has several implications for how different representatives affect the acceptance set. Extremist representatives induce larger acceptance sets than their aligned moderates, with these sets more extreme on both ends. For instance, any left-moderate representative produces an acceptance set strictly contained in $[\underline{x}_\ell, \overline{x}_\ell]$. The two extremist acceptance sets can be ordered: the most leftward proposals occur when d is a left-extremist, while the most rightward proposals occur when d is a right-extremist. Finally, moderate representatives can shift the set of proposals that pass even without becoming the median politician, highlighting that the principal's choice matters not just for who is pivotal but also for the strategic environment other politicians face.

Lemma 3 also establishes smoothness properties of the acceptance set boundaries. The functions $\underline{x}(y_d)$ and $\overline{x}(y_d)$ are differentiable almost everywhere on $(\underline{x}_\ell, \overline{x}_r)$. Kinks occur at two types of locations. First, at y_ℓ and y_r, marking transitions from moderate to centrist, the derivative jumps discontinuously because the acceptance set shifts more sharply over centrists than over moderates, as Figure 3 illustrates. Second, kinks occur at any y_d where some politician $i \in K$ satisfies $y_i = \underline{x}(y_d)$ or $y_i = \overline{x}(y_d)$—where i's ideal point coincides with an acceptance set boundary, causing i to switch between proposing y_i and proposing that boundary. On intervals of smoothness, $\underline{x}(y_d)$ is strictly concave and $\overline{x}(y_d)$ strictly convex, due to quadratic utility.

6 Optimal Representation

We have characterized how the representative's ideal point affects the acceptance set and thereby the proposals made by all politicians. We now characterize which representatives are optimal for different principals.

Since each y_d induces a unique policy lottery (Lemma 1), the principal's expected utility given $y_p \in X$ is uniquely defined over $y_d \in X$ as $U(y_d, y_p) \equiv V_p(A(y_d))$. The principal's problem is to choose y_d to maximize $U(y_d, y_p)$. We characterize the optimal representative correspondence $y_d^* : X \rightrightarrows X$, where:

$$y_d^*(y_p) \equiv \underset{y_d \in X}{\operatorname{argmax}}\, U(y_d, y_p) \tag{4}$$

is non-empty, upper hemicontinuous, and compact-valued since $U(y_d, y_p)$ is continuous in y_d.

This is a complex optimization problem. The lotteries induced by different representatives need not be ordered in any simple way, as shifts in y_d simultaneously affect which politicians are constrained, where they propose, and the acceptance set itself. Different representatives could generate equivalent outcomes for the principal by inducing different policymaking that trades off the representative's own proposal against extremist proposals in offsetting ways. Since y_d affects both the mean and variance of policy outcomes through multiple channels, standard single-crossing arguments need not apply globally.

Despite this complexity, Proposition 1 reveals a striking pattern. The optimal representative correspondence is monotone increasing: more right-leaning principals choose more right-leaning representatives. A wide interval of principals strictly bias their representative toward the center, and no principal biases away from center. The principal's location determines which type of representative she prefers: very extreme principals are potentially indifferent among all extremists from their side, less extreme principals choose a unique moderate from their side who is more centrist than themselves, and centrist principals choose centrists.

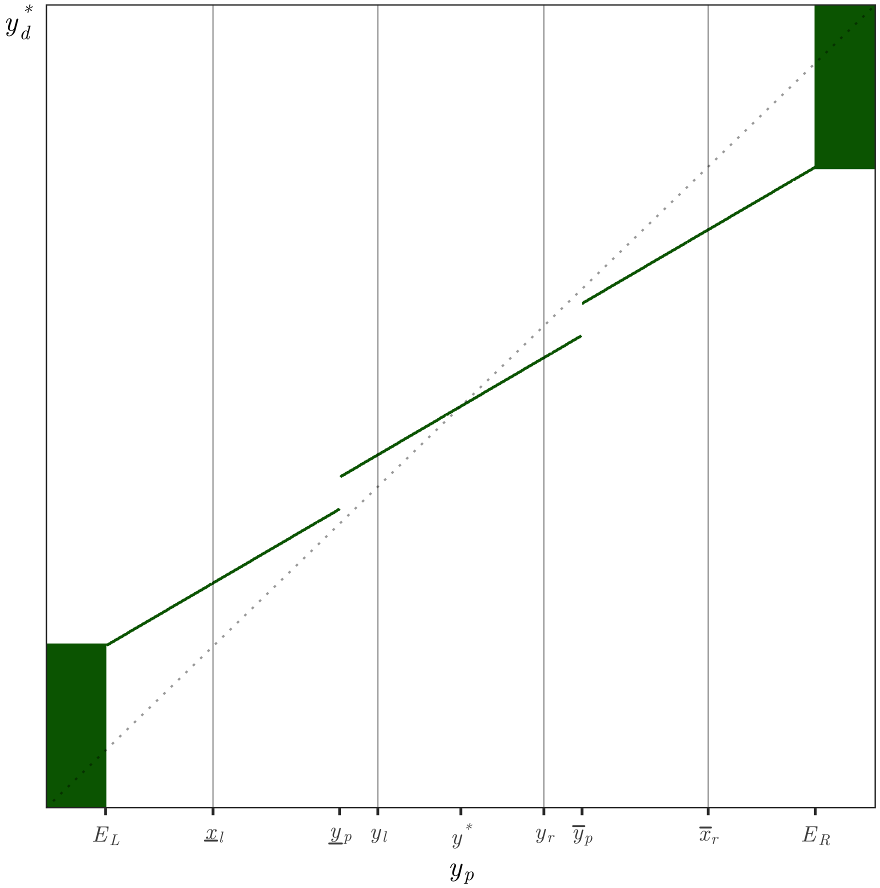
Proposition 1. *The optimal representative correspondence y_d^* is increasing in the strong set order sense. Moreover, there are intervals (E_L, E_R) and $(\underline{y}_p, \overline{y}_p)$ satisfying $(E_L, E_R) \supset (\underline{x}_\ell, \overline{x}_r) \supset [\underline{y}_p, \overline{y}_p] \supset [y_\ell, y_r]$ such that y_d^* is single-valued and continuous almost everywhere on (E_L, E_R) and:*

1. *$y_p < E_L$ implies $y_d^*(y_p) = [0, \underline{x}_\ell]$;*
2. *$y_p \in (E_L, \underline{y}_p)$ implies $y_d^*(y_p) \subset (y_p, y_\ell)$;*
3. *$y_p \in (\underline{y}_p, \overline{y}_p)$ implies $y_d^*(y_p) \subset [y_\ell, y_r]$;*
4. *$y_p \in (\overline{y}_p, E_R)$ implies $y_d^*(y_p) \subset (y_r, y_p)$; and*
5. *$y_p > E_R$ implies $y_d^*(y_p) = [\overline{x}_r, 1]$.*

Proposition 1 reveals that almost all principals strictly bias their representative toward the center, with the strength and nature of this bias depending on the principal's ideological position. To understand why, we establish which types of representatives could be optimal and how the principal's tradeoffs vary with her location. Figure 4 illustrates the main patterns.

Non-centrists never choose representatives from the other side. No principal with $y_p \leq y_r$ finds it optimal to choose $y_d > y_r$, and vice versa. A representative from the opposite side shifts both their own proposal and

Figure 4: Optimal representatives



Note: Figure 4 illustrates Proposition 1 for a five-member legislature where $\delta = .99$, $(\rho_1, ..., \rho_4, \rho_d) = (.2, .25, .25, .2, .1)$, and $(y_1, ..., y_4) = (0, .4, .6, 1)$. Very extreme principals ($y_p \notin (E_L, E_R)$) are indifferent among all extremist representatives from their side. Intermediate principals ($y_p \in (E_L, \underline{y}_p) \cup (\overline{y}_p, E_R)$) choose a unique moderate representative from their side who is strictly more centrist than themselves. Centrist principals ($y_p \in (\underline{y}_p, \overline{y}_p)$) choose a unique centrist representative. Almost all principals shift their representative toward the center relative to their own position.

extremist proposals away from the principal. While this might affect the nearest acceptance set boundary, moving y_d toward the principal's side achieves the same (or better) effect on that boundary while avoiding the cost of pushing the far boundary further away. For a principal with $y_p \leq y_r$, shifting y_d to the right of y_r pushes both d's proposal and right-extremist proposals away from y_p, so any benefit must come from pulling $\underline{x}(y_d)$ closer to y_p. But by Lemma 3, shifting y_d in either direction away from y_r pulls $\underline{x}(y_d)$ leftward. Thus any principal who benefits from shifting y_d to the right of y_r can achieve the same benefit by instead shifting y_d leftward without the cost of shifting $\overline{x}(y_d)$ further away.

Extremist and moderate principals choose more centrist representatives. For moderate principals, choosing an extremist from their side provides no benefit—it only pushes proposals further away. Choosing a more centrist representative, however, constrains extremist proposals on both sides. Consider a moderate principal at $y_p \in (\underline{x}_\ell, y_\ell)$. If she chooses $y_d = y_p$, then $y_p \in A(y_d)$ so d would propose y_p if recognized. Shifting y_d leftward pushes both d's proposal and extremist proposals further from y_p, so $y_d < y_p$ is never optimal. Shifting y_d rightward toward y_ℓ pulls both boundaries of $A(y_d)$ closer to y_p at the cost of pushing d's proposal further away. At $y_d = y_p$, the marginal cost of this tradeoff is zero (since d already proposes y_p), while the marginal benefit from constraining extremists is strictly positive. Thus $y_p < \min y_d^*(y_p)$ for all moderate principals.

Extremist principals choose either an aligned extremist or an aligned moderate, but never a centrist. For a left-extremist principal, centrist representatives induce lotteries that first-order stochastically dominate those induced by left-extremists (since $A(y_d)$ shifts entirely rightward over centrists), making centrists strictly worse. However, extremist principals may choose moderates over extremists. Starting at $y_d = \underline{x}_\ell$ and shifting rightward into the left-moderate region pushes d's proposal and aligned extremist proposals away from y_p but pulls opposing extremist proposals closer. If y_p is sufficiently close to $\underline{x}_\ell$, the marginal cost of constraining opposing extremists is near zero, so there exists a non-empty interval of extremists who choose an aligned moderate. Upper hemicontinuity of y_d^* implies that there exists a unique threshold $E_L \in (0, \underline{x}_\ell)$ such that $y_d^*(E_L) = [0, \underline{x}_\ell]$ (principals are indifferent between all aligned extremists), and $\underline{x}_\ell < \min y_d^*(y_p)$ for all $y_p \in (E_L, \underline{x}_\ell)$. If no left-extremist finds moderates optimal, then $E_L = 0$. The threshold E_R is defined analogously.

The ideological ordering is preserved. More right-leaning principals choose weakly more right-leaning representatives. The conceptual intuition has two parts. First, over the set of representatives actually chosen by some principal, more rightward representatives produce policy lotteries with higher average outcomes. Second, principals prefer higher average policy outcomes when they themselves are more right-leaning. Together, these imply monotonicity.

The first step is not automatic. In principle, a more moderate representative could produce higher average outcomes than a more extreme representative on the same side, due to how the acceptance set contracts as the representative moves toward the median. However, when this pattern occurs, shifting the representative in that direction both moves the average outcome closer to principals on that side and reduces uncertainty in outcomes. Both effects are beneficial for those principals, so such representatives are never optimal. Instead, principals always prefer to continue shifting further. Thus, over the set of representatives actually chosen, more rightward representatives produce higher average outcomes.

Establishing this monotonicity requires showing that, over the set of representatives actually chosen by some principal, more rightward representatives produce outcomes that more right-leaning principals prefer. The key step is to show that the mean policy $\mu(y_d)$ is weakly increasing over this set. Since preferences are quadratic, any player's utility reduces to $U(y_d, y_p) = 1 - (y_p - \mu(y_d))^2 - \sigma^2(y_d)$, where $\mu(y_d)$ and $\sigma^2(y_d)$ are the mean and variance of the induced policy lottery. If $\mu(y_d)$ is increasing, then U satisfies single-crossing and monotonicity of y_d^* follows.

Over centrists, monotonicity of $\mu(y_d)$ is immediate: both boundaries of $A(y_d)$ strictly increase with y_d (Lemma 3), so lotteries are ordered by first-order stochastic dominance and $\mu(y_d)$ is strictly increasing. The challenge is that over some moderates, $\mu(y_d)$ can potentially decrease if extremist proposal rights are highly asymmetric. As a left-moderate representative shifts toward y_ℓ, the acceptance set contracts with both boundaries moving inward at equal rates. The lower bound and d's proposal shift rightward (raising μ), while the upper bound shifts leftward (lowering μ). If right-extremists hold sufficient proposal rights, the contraction effect can dominate locally.

However, we can rule out such representatives being optimal for any principal. On any interval where $\mu(y_d)$ decreases, $\sigma^2(y_d)$ also decreases and the mean remains above y_ℓ (bounded away from the median). For all principals with $y_p < y_\ell$, shifting y_d rightward in such a region both pulls $\mu(y_d)$ closer to y_p and reduces variance, making $U(y_d, y_p)$ strictly increasing. Thus no principal chooses representatives from such intervals.

It follows that $\mu(y_d)$ is weakly increasing over the image of y_d^*, establishing that y_d^* is increasing in the strong set order sense.

Lemma 4 establishes these observations about when $\mu(y_d)$ is monotone and characterizes key properties of how the policy lottery varies with y_d.

Lemma 4. *There exist $\underline{\pi} \in [\underline{x}_\ell, y_\ell)$ and $\overline{\pi} \in (y_r, \overline{x}_r]$ such that $\mu(y_d)$ is strictly increasing on $[\underline{\pi}, \overline{\pi}]$. Moreover, $\underline{\pi} > \underline{x}_\ell$ implies $1 - \tilde{P}(y_\ell) > \frac{1}{2\delta}$ and $\overline{\pi} = \overline{x}_r$; and $\overline{\pi} < \overline{x}_r$ implies $\tilde{P}(y_\ell) > \frac{1}{2\delta}$ and $\underline{\pi} = \underline{x}_\ell$. Additionally, on any interval $Z \subset [\underline{x}_\ell, \underline{\pi}]$ such that $\mu(y_d)$ is decreasing, $y_\ell < \mu(y_d)$ and $\sigma^2(y_d)$ is strictly decreasing; and on any interval $Z \subset [\overline{\pi}, \overline{x}_r]$ such that $\mu(y_d)$ is decreasing, $\mu(y_d) < y_r$ and $\sigma^2(y_d)$ is strictly increasing.*

Almost all non-extremist principals have unique optimal representatives. The optimal representative correspondence y_d^* is single-valued almost everywhere on (E_L, E_R). This uniqueness follows from the strict monotonicity of $\mu(y_d)$ over representatives actually chosen. Since $\mu(y_d)$ is strictly increasing on the image of y_d^* restricted to (E_L, E_R) and preferences satisfy strict single-crossing in y_d when $\mu(y_d)$ is strictly increasing, any selection from $y_d^*|_{(E_L,E_R)}$ must be strictly increasing. Consequently, y_d^* can be multi-valued only at isolated points where the principal's payoff $U(y_d, y_p)$ has kinks—specifically, at the boundaries between moderate and centrist regions ($\underline{y}_p$ and $\overline{y}_p$) and potentially at other points where politicians enter or exit the acceptance set.

Some representatives are optimal for multiple principals, while others are not optimal for anyone. There are *dead zones*—open intervals of representatives who are not optimal for any principal. There is always at least one around a boundary between moderate and centrist regions (i.e., around y_ℓ or y_r). Moderate principals want to constrain extremists by choosing more centrist representatives, but as a moderate representative approaches the median (say $y_d \to y_\ell$), the marginal benefit from further constraining extremists diminishes to zero while the cost of pushing the representative's proposal away remains positive. At some point, it becomes optimal to jump discontinuously over the boundary and choose a strict centrist rather than a moderate arbitrarily close to y_ℓ.

More precisely, the boundaries $\underline{y}_p$ and $\overline{y}_p$ separate principals who prefer moderates from those who prefer centrists. As moderate representatives approach either boundary, the acceptance set changes increasingly slowly—by Lemma 3, both boundaries of $A(y_d)$ change negligibly as $y_d \to y_\ell$ or $y_d \to y_r$. For moderates sufficiently close to y_ℓ or y_r, choosing a representative at this boundary provides minimal benefit in constraining extremists while still pushing the representative's proposal away from the principal. Such principals may prefer to jump over the boundary and choose a strict centrist. This creates discontinuities in y_d^* at $\underline{y}_p$ and $\overline{y}_p$, where the correspondence may contain both a moderate and a centrist.

More broadly, y_d^* can have multiple discontinuities since there can be dead zones around other kinks in $U(y_d, y_p)$ (i.e., where some politician $i \in K$ enters or exits the acceptance set).

Self-representation. All sufficiently extreme principals weakly prefer to self-represent since they are indifferent among all extremists on their side. Among non-extremist principals in (E_L, E_R), only a finite set of centrists strictly prefer self-representation. All other non-extremist principals strictly prefer representatives more centrist than themselves.

The logic differs across types. For extremist principals with $y_p < E_L$, every aligned extremist representative induces the same acceptance set $[\underline{x}_\ell, \overline{x}_\ell]$ (by Lemma 3), so the principal is indifferent among all such representatives, including herself. For moderate principals, self-representation is never optimal since they always strictly prefer more centrist representatives. For centrist principals, the calculus is more subtle. A centrist representative is the median, so shifting y_d away from y_p will shift both boundaries in the same direction. The principal's marginal gain from shifting y_d rightward depends on the balance between constraining left-extremists (who propose $\underline{x}$) and right-extremists (who propose $\overline{x}$). Only for special centrist positions—where the marginal benefit of constraining each side exactly balances—is self-representation optimal. For other centrist positions, the principal strictly prefers to bias her representative toward one side.

Corollary 1 formalizes this observation and identifies the focal representatives.

Corollary 1 (Focal Representatives)**.** *There exists a non-empty finite set* $Y^* \subset [y_\ell, y_r]$ *such that if* $y_p \in (E_L, E_R)$*, then* $y_p \in y_d^*(y_p)$ *if and only if* $y_p \in Y^*$*. Moreover,* $y_d^* \subset (y_p, \min Y^*)$ *for all* $y_p \in (E_L, \min Y^*)$ *and* $y_d^* \subset (\max Y^*, y_p)$ *for all* $y_p \in (\max Y^*, E_R)$*.*

Among the finite set Y^* of centrists who can self-represent, there are two extreme focal points: the leftmost, $y_d = \min Y^*$, serves as a locus of attraction for all left-leaning non-extremist principals, and the rightmost, $y_d = \max Y^*$, attracts all right-leaning non-extremist principals. All principals in (E_L, E_R) who do not self-represent strictly shift their representative toward the center—specifically, toward one of these extreme focal points.

6.1 Polarized Legislature

To illustrate our general results and sharpen our characterizations, we now analyze a tractable special case.

Assumption 2 (Polarized Legislature)**.** *There are two politicians,* ℓ *and* r*, whose ideal points satisfy* $y_\ell, y_r \in \text{int } A(y_d)$ *for all* $y_d \in X$*, while all other politicians in* K *are always constrained* ($y_i \notin \text{int } A(y_d)$ *for all* $y_d \in X$*).*

This assumption captures settings such as a two-party legislature where each party contains moderate and extreme factions, with the moderates close enough together that they always lie in the interior of $A(y_d)$ regardless of who serves as representative. It is satisfied by a legislature with two extreme factions at $y_L = 0$ and $y_R = 1$, and two moderate factions at positions y_ℓ and y_r that are sufficiently close to each other. Specifically, $(y_r - y_\ell) < (1-\delta)/2$, ensures that $y_\ell, y_r \in \text{int } A(y_d)$ for all y_d under our high patience assumption.

The key advantage of this structure is that *which* politicians are constrained no longer varies with y_d. In the general model, shifting the representative's position can cause individual politicians to move in or out of the acceptance set, creating kinks in the acceptance set boundaries. Under Assumption 2, the set of constrained politicians is fixed—only the extremists are ever constrained—so the acceptance set boundaries $\underline{x}(y_d)$ and $\overline{x}(y_d)$ are continuously differentiable over each ideological region (left-extremists, left-moderates, centrists, right-moderates, right-extremists). The only kinks occur at the boundaries between these regions: at y_ℓ and y_r, where the representative transitions from moderate to centrist and the acceptance set shifts more sharply, as depicted in Figure 3. Additionally, on smooth intervals, $\underline{x}(y_d)$ is strictly concave and $\overline{x}(y_d)$ is strictly convex due to quadratic utility.

This smoothness sharpens the characterization of optimal representatives. Assumption 2 clarifies intuitions from Proposition 1 and establishes several stronger results. Let ρ_L denote the aggregate proposal rights of politicians with $y_i < y_\ell$ (left-extremists) and ρ_R denote the aggregate proposal rights of politicians with $y_i > y_r$ (right-extremists).

(i) Why optimal representatives are monotone increasing. Proposition 1 established that more right-leaning principals choose weakly more right-leaning representatives. Under Assumption 2, the underlying logic becomes clearer and the result stronger. The mean policy $\mu(y_d)$ is strictly increasing over the entire

interval of non-extremist representatives $(\underline{x}_\ell, \overline{x}_r)$, not just over the subset of representatives actually chosen by some principal.

In the general case, $\mu(y_d)$ can decrease over some moderate representatives when extremist proposal rights are highly asymmetric. We established monotonicity of y_d^* by showing that such representatives—where μ decreases—are never optimal for any principal. Under Assumption 2, this exclusion is unnecessary: $\mu(y_d)$ never decreases in the first place. The set of constrained proposers is fixed across all y_d, so the acceptance set boundaries vary smoothly (except at y_ℓ and y_r). This rules out the complex interactions between the representative's proposal, the contraction of the acceptance set, and changing extremist weights that could produce local decreases in $\mu(y_d)$ in the general case. Over centrists, $\mu(y_d)$ increases because both boundaries shift rightward. Over moderates, the acceptance set contracts, but the smoothness and fixed set of constrained proposers ensures that the mean strictly increases. This delivers strict single-crossing of $U(y_d, y_p)$ over the entire range $(\underline{x}_\ell, \overline{x}_r)$ and guarantees the strict monotonicity visible in Figure 4.

(ii) There is a unique locus of attraction. In general, the set of principals who strictly prefer self-representation is a finite subset $Y^* \subset [y_\ell, y_r]$ of centrists. Under Assumption 2, this set is a singleton: there exists a unique focal representative y^* toward which all non-extremist principals bias their choices.

Remark 2 (Unique Focal Representative)**.** *Under Assumption 2, there exists a unique* $y^* \in [y_\ell, y_r]$ *such that* $y_d^*(y_p) = \{y^*\}$ *if* $y_p = y^*$*,* $y_d^*(y_p) \subset (y_p, y^*)$ *for all* $y_p \in (E_L, y^*)$*, and* $y_d^*(y_p) \subset (y^*, y_p)$ *for all* $y_p \in (y^*, E_R)$*.*

Consider a centrist principal at $y_p \in [y_\ell, y_r]$. At $y_d = y_p$, the marginal effect of shifting y_d on the representative's own proposal is zero (since d already proposes y_p), so the marginal gain comes entirely from how the acceptance set boundaries move. Since this representative would be the median, the principal's marginal gain from shifting y_d rightward is:

$$\lambda(y_p) \equiv \rho_L \frac{\partial \underline{x}(y_d)}{\partial y_d}\bigg|_{y_d = y_p} - \rho_R \frac{\partial \overline{x}(y_d)}{\partial y_d}\bigg|_{y_d = y_p}. \tag{5}$$

This expression captures the tradeoff: shifting right pulls in the left boundary (benefiting the principal with weight ρ_L, the proposal rights of left-extremists) but pushes out the right boundary (hurting the principal with weight ρ_R). Since $\underline{x}$ is strictly concave and $\overline{x}$ is strictly convex over $[y_\ell, y_r]$, λ is strictly decreasing in y_p. Thus λ crosses zero exactly once, defining the unique y^* where a centrist principal is indifferent between biasing left or right. All centrists with $y_p < y^*$ have $\lambda(y_p) > 0$ and prefer biasing rightward toward y^*, while all with $y_p > y^*$ prefer biasing leftward toward y^*.

(iii) Complete characterization of dead zones. Proposition 1 established that dead zones exist around at least one of y_ℓ or y_r. Under Assumption 2, we can fully characterize when and where these dead zones occur. The key is that dead zones arise when moderates near a boundary prefer to jump over it to choose centrists, which depends on whether the unique focal representative y^* lies at that boundary or in the interior.

If $y^* \in (y_\ell, y_r)$, then dead zones exist around *both* y_ℓ and y_r: moderates on both sides want to jump toward the interior focal point y^*, so representatives near either boundary are suboptimal. If $y^* = y_\ell$, a dead zone exists only around y_r: left-moderates find representatives near y_ℓ optimal (the focal point lies at the boundary), but right-moderates want to jump leftward toward y^*, creating a dead zone around y_r. Symmetrically, if $y^* = y_r$, then a dead zone exists only around y_ℓ.

Where is y^* located? From equation (5), the location of y^* depends on the balance between ρ_L and ρ_R. The sign of $\lambda(y_\ell)$ determines whether $\underline{y}_p < y_\ell$ or $\underline{y}_p = y_\ell$, and similarly the sign of $\lambda(y_r)$ determines whether $y_r < \overline{y}_p$ or $\overline{y}_p = y_r$. If $\lambda(y_r) < 0 < \lambda(y_\ell)$, then moderates near both boundaries want to bias inward, so

$y^* \in (y_\ell, y_r)$ and dead zones exist on both sides. If $\lambda(y_\ell) \leq 0$, then $y^* = y_\ell$ and only right-moderates create a dead zone. If $\lambda(y_r) \geq 0$, then $y^* = y_r$ and only left-moderates create a dead zone. Notably, representatives at y_ℓ or y_r can be optimal for: (i) nobody, (ii) exactly one principal, or (iii) an interval of centrist principals. They are the only representatives who can be uniquely optimal for more than one principal.

(iv) Comparative statics in extremist power. Assumption 2 permits clean comparative statics on the balance of extremist proposal power. Fix total extremist power $\rho_E = \rho_L + \rho_R$ and consider transferring agenda-setting power from right-extremists to left-extremists (increasing ρ_L at the expense of ρ_R). This shift does not change the acceptance set, since the median remains indifferent between the two extremist groups' proposals, but it does change the principal's delegation incentives through three effects.

First, the locus of attraction shifts away from the empowered side. From equation (5), increasing ρ_L raises the weight on $\partial \underline{x}/\partial y_d$ in $\lambda(y_p)$. Since λ is strictly decreasing in y_p, its unique zero shifts rightward. Centrist principals become more concerned about constraining the empowered left-extremists and more of them want to bias their representative rightward.

Second, the set of principals who choose non-extremist representatives expands on the weakened side and contracts on the strengthened side. Specifically, the thresholds E_L and E_R both shift rightward. Extremist principals want to constrain opposing extremists but not aligned extremists. On the strengthened left side, moderating becomes less appealing because constraining opposing (right) extremists yields lower returns and constraining aligned (left) extremists is more costly. On the weakened right side, the calculus is flipped, so moderating becomes more appealing due to greater returns from constraining left-extremists. Thus more right-extremists choose moderates while fewer left-extremists do.

Third, the locations of dead zones can also change. Since the location of y^* depends on the balance between ρ_L and ρ_R, changes in extremist power determine whether dead zones exist around both boundaries, only the left, or only the right. Empowering the left shifts y^* rightward, potentially moving it from the interior (y_ℓ, y_r) to the right boundary y_r, or from the left boundary y_ℓ into the interior. This affects which representatives are not optimal for any principal.

6.2 Competitive Representation

Thus far, we have focused on a single principal filling one position with the rest of the legislature fixed. This setting naturally captures situations where other politicians are already in office, such as executive appointments to existing legislative bodies or committee assignments. Our analysis suggests that incentives for strategic moderation will arise when multiple positions are filled simultaneously, as noted by Gailmard and Hammond (2011). We now explore whether those incentives strengthen or weaken by extending the model to allow multiple principals to simultaneously choose their representatives.

We begin with a two-principal case to illustrate key strategic forces. We then show that the qualitative pattern—strategic moderation toward a common locus of attraction—extends to the general case of n principals, though with key differences in which representative serves as the locus of attraction.

6.2.1 The Two-Principal Case

Consider principals P_a and P_b with ideal points $y_{p_a} < y_m < y_{p_b}$, each simultaneously appointing representatives a and b to fill two positions in a five-player body with a fixed median m and two extremists L and R. We assume y_{p_a} and y_{p_b} are always inside the acceptance set while y_L and y_R are always outside. We use (y_a^*, y_b^*) to denote the pair of representatives chosen in equilibrium.

Since both principals are moderates relative to the fixed median, each biases their representative toward m in equilibrium. With quadratic utility, each principal has a unique best response. Principal P_a's best response to y_b, denoted $y_a(y_b)$, is characterized by the first-order condition:

$$\frac{\partial \underline{x}(y_a, y_b)}{\partial y_a}\left(\rho_L(y_{p_a} - \underline{x}(y_a, y_b)) + \rho_R(\overline{x}(y_a, y_b) - y_{p_a})\right) - \rho_a(y_a - y_{p_a}) = 0, \tag{6}$$

with P_b's best response function analogous.

The key comparative static concerns how each principal responds to moderation by their opponent. When P_b shifts y_b inward (toward y_m), this creates two countervailing effects on P_a's incentives. A *substitution effect*: extremist proposals shift inward, directly benefiting P_a and reducing her marginal benefit from her own moderation. Through this channel, moderation by P_b *substitutes* for moderation by P_a. A *complementarity effect*: the acceptance set becomes more responsive to y_a (specifically, $\frac{\partial^2 \underline{x}(y_a, y_b)}{\partial y_a \partial y_b} < 0$), reducing the "price" of moderating extremist proposals. Through this channel, moderation by P_b *complements* moderation by P_a.

Which effect dominates depends on the balance of extremist proposal rights.

Lemma 5. *If $\rho_L < \rho_R$, then $y_a(y_b)$ is strictly decreasing in y_b and $y_b(y_a)$ is strictly increasing in y_a. If $\rho_L > \rho_R$, these relationships reverse. If $\rho_L = \rho_R$, then $y_{d_i}(y_{d_{-i}}) = (1 - \delta\rho_E)y_{p_i} + \delta\rho_E y_m$ for all y_{-i}.*

The monotonicity of best responses ensures a unique equilibrium exists, with both principals choosing strictly interior positions:

Remark 3. *There is a unique equilibrium, in which $y_a^* \in (y_{p_a}, y_m)$ and $y_b^* \in (y_m, y_{p_b})$.*

Comparing equilibrium choices to the single-principal benchmark reveals an asymmetry based on extremist power. When extremist power is balanced ($\rho_L = \rho_R$), the increased elasticity of the acceptance set exactly offsets the decreased marginal benefit of moderation, leaving each principal's choice unchanged. As one side's extremists gain power, the complementarity effect dominates for the principal on the weak side (who moderates more than in the single-principal case), while the substitution effect dominates for the principal on the strong side (who moderates less).

Corollary 2. *Let $y^s{*}_i$ denote principal i's optimal choice in the single-principal model, fixing $y_{-i} = y_{p_{-i}}$. In the two-principal setting, we have the following in equilibrium: (i) $\rho_L < \rho_R$ implies $y_a^* < y^s{*}_a < y_m < y_b^* < y^s{*}_b$; (ii) $\rho_L = \rho_R$ implies $y_a^* = y^s{*}_a < y_m < y_b^s = y^*{*}_b$; and (iii) $\rho_L > \rho_R$ implies $y^s{*}_a < y_a^* < y_m < y^s{*}_b < y_b^*$.*

6.2.2 General Case: Multiple Principals

The two-principal case reveals that competition introduces strategic interactions between principals' moderation choices, with the direction of these interactions depending on the balance of non-appointee power. We now show that the qualitative pattern of strategic moderation persists with n principals, though the locus of attraction differs qualitatively from the two-principal case.

We consider n (odd) principals with ideal points $y_1 < y_2 < ... < y_n$, where all ideal points are common knowledge. All principals simultaneously choose the ideology of their representative, $d_i \in X$. The n representatives then bargain over policy as in the baseline model, with each representative receiving recognition probability ρ_i. Let $d = (d_1, d_2, \ldots, d_n)$ denote the profile of representative ideal points and d_{-i} denote the profile excluding d_i. All principals are purely policy-motivated, as in the baseline model.

An *equilibrium* consists of (i) a profile of stationary bargaining strategies in the bargaining subgame and (ii) a profile of representative ideal points d^* in which d_i^* is a best response to d_{-i}^* for all i. Throughout this section, we focus on equilibria in which any extremist principal whose ideal point lies outside the equilibrium acceptance set selects the nearest boundary representative rather than a more extreme choice (among which they are indifferent). Given any profile d_{-i}, let ℓ_i and r_i denote the $\frac{n-1}{2}$-th and $(\frac{n-1}{2}+1)$-th representative ideal points in the profile d_{-i}.

Proposition 1 directly characterizes principal i's best response to any d_{-i}: given the other representatives d_{-i}, principal i chooses d_i exactly as in the single-principal model, treating d_{-i} as the exogenous legislature. A key difference from the two-principal case is that the median principal's representative is the unique locus of attraction toward which all other principals bias their representative. Let $m \equiv \frac{n+1}{2}$ denote the median principal.

Proposition 2 (Competitive Representation with Multiple Principals)**.** *In any equilibrium:*

1. *The median principal's representative is the median politician: d_m^* is the median of $(d_1^*, ..., d_n^*)$.*
2. *The ideological ordering of principals is preserved: $d_i^* \leq d_m^*$ for all $i < m$ and $d_i^* \geq d_m^*$ for all $i > m$.*
3. *All non-median principals shift toward the median principal's representative: $d_i^* \in (y_i, d_m^*]$ for all $i < m$ and $d_i^* \in [d_m^*, y_i)$ for all $i > m$.*

Consequently, d_m^* serves as a unique locus of attraction: all principals choose representatives strictly between their own ideal point and d_m^*.

Proposition 2 follows largely from our earlier characterizations and has three key insights. First, only the median principal's representative can be the median politician: any principal i's representative is the median politician only if i would be the median principal under self-representation.

Second, the ideological ordering of representatives matches the ordering of their principals. If some non-median principal $i < m$ chose $d_i > d_m$, then from principal i's perspective the median would be at or left of d_m, implying $\ell_i = d_m$. But Proposition 1 establishes that principals only choose representatives from their own side of the spectrum, contradicting $d_i > d_m \geq y_m > y_i$.

Third, there is strategic moderation. For any non-median principal, the median principal's representative d_m^* serves the same role as y_ℓ or y_r in the baseline model—it becomes the boundary between centrist and moderate regions. Since our baseline analysis established that moderate principals strictly prefer representatives more centrist than themselves (choosing from (y_i, y_ℓ) if left-leaning), non-median principals must choose d_i strictly between their own position and d_m^*.

Proposition 2 has three implications. First, it shows that *strategic moderation is robust to competition*: even when all principals simultaneously optimize, moderation persists and follows a predictable pattern. Second, unlike the single-principal model where the locus of attraction y^* depends on the balance of extremist power in the fixed legislature, here the locus corresponds to whichever representative the median principal selects. This differs from the two-principal case, where both principals moderate toward the fixed median m and the balance of extremist power affects the degree of moderation but not the target. Third, the result suggests testable predictions: in any institutional setting where multiple principals appoint representatives to a collective body, we should observe (i) the median appointee representing the median principal and (ii) all other appointees positioned strictly between their principal and the median appointee.

The equilibrium may not be unique. Different equilibria can feature different values of d_m^*, with the pattern of strategic moderation adjusting accordingly. Within any given equilibrium, however, the locus of attraction is unique and corresponds to the median principal's choice, so all principals shift toward whichever median representative emerges.

7 Conclusion

We study how principals choose representatives who participate in collective policymaking and how these choices shape policy outcomes. A key force in our analysis is that a representative's ideology affects legislature-wide expectations about policymaking. We show how this generates *anticipation effects*, shaping which policies each politician will support and thereby influencing what can pass and what extremists propose.

We provide a general logic for why moderate representatives appeal to principals. All centrist principals want a centrist representative who will be the median (de facto veto) politician, while all moderate principals want a more centrist representative. Even when not the median, a more centrist representative improves the median's expectations about proposals, narrowing what can pass and constraining extremist politicians.

We focus on collective policymaking governed by simple majority rule. Our results generalize to any *strong* voting rule, since a single decisive politician effectively determines what can pass (Duggan 2014). Several natural extensions emerge. Settings where the median is effectively fixed isolate pure anticipation effects—where representatives shape outcomes solely through bargaining expectations rather than pivotal voting. Such settings arise when appointing proposers to bargain with fixed veto players. Welfare gains from strategic delegation peak for principals at ideological boundaries where tradeoffs between constraining extremists and maintaining favorable proposals are sharpest. Understanding how institutional features concentrate or disperse these gains could indicate where delegation is most consequential.

Studying representation under supermajority rules, where multiple politicians are simultaneously decisive, remains an open question. Under such rules, the acceptance set is determined by two endogenous veto players rather than a single median, adding considerable complexity. Whether supermajority requirements strengthen or weaken moderation incentives would deepen our understanding of delegation across institutions.

Our results have direct implications for representation in separation-of-powers systems and congressional committees. The framework also suggests new insights for studying elections to collective bodies, including voter behavior (Kedar 2005, 2009; Duch et al. 2010), electoral competition (Austen-Smith and Banks 1988; Krasa and Polborn 2018), and representativeness (Austen-Smith and Banks 1991). When groups collectively choose representatives, mild conditions ensure standard spatial voting results apply—the median principal becomes decisive and coalitions are spatially ordered—suggesting our logic extends naturally to electoral settings. We shed new light on how expectations about collective policymaking affect incentives of party leaders and voters, influencing who gets nominated and their electoral prospects. A key complication is that the policymaking environment affects not only how parties evaluate their own candidates but also how they view opposing candidates. Understanding how elite polarization and extremism shape electoral competition for positions in collective bodies remains a rich area for future research.

Appendix

A Equilibrium Policymaking

We use the following in our proofs to economize on notation.

Definition 2. *$M_L = [\underline{x}_\ell, y_\ell]$, $M_R = [y_r, \overline{x}_r]$, and $M = M_L \cup M_R$.*

A.1 Proof of Lemma 1.

Follows from Propositions 1–2 in Cardona and Ponsati (2011). $\square$

A.2 Proof of Lemma 2.

From Cardona and Ponsati (2011), the equation

$$u(x, y_\ell)(1 - \delta\rho_d) - \delta\left(\sum_{i\in K:|y_i - y_\ell|\geq|y_\ell - x|} \rho_i u(x, y_\ell) + \sum_{i\in K:|y_i - y_\ell|<|y_\ell - x|} \rho_i u(y_i, y_\ell)\right) = 0 \tag{7}$$

has exactly two solutions: $\underline{x}_\ell \in (0, y_\ell)$ and $\overline{x}_\ell \in (y_\ell, 1)$ where solutions are interior due to Assumption 1.

Similarly, $\underline{x}_r \in (0, y_r)$ and $\overline{x}_r \in (y_r, 1)$ are the only solutions of

$$u(x, y_r)(1 - \delta\rho_d) - \delta\left(\sum_{i\in K:|y_i - y_r|\geq|y_r - x|} \rho_i u(x, y_r) + \sum_{i\in K:|y_i - y_r|<|y_r - x|} \rho_i u(y_i, y_r)\right) = 0. \tag{8}$$

Part 1 of the proof shows that $y_d \in (\underline{x}_\ell, \overline{x}_r)$ implies $y_d \in \text{int } A(y_d)$. Part 2 shows that $y_d \in \text{int } A(y_d)$ implies $y_d \in (\underline{x}_\ell, \overline{x}_r)$. To show each direction, we use contraposition.

Part 1. Consider $y_d \leq \min A(y_d) = \underline{x}(y_d)$. Then $y_d \leq \underline{x}(y_d) < y_m$, so Assumption 1 implies that $\underline{x}(y_d) \in (0, y_\ell)$ and must solve (7). Thus, $\underline{x}(y_d) = \underline{x}_\ell$. Analogously using (8), $y_d \geq \max A(y_d) = \overline{x}(y_d)$ implies that $\overline{x}(y_d) = \overline{x}_r$. We have shown that $y_d \notin (\underline{x}(y_d), \overline{x}(y_d))$ implies $y_d \notin (\underline{x}_\ell, \overline{x}_r)$. By contraposition, $y_d \in (\underline{x}_\ell, \overline{x}_r)$ implies $y_d \in (\underline{x}(y_d), \overline{x}(y_d)) = \text{int } A(y_d)$.

Part 2. Consider $y_d \leq \underline{x}_\ell$. Then, uniqueness of $A(y_d)$ implies that $y_d \notin \text{int } A(y_d)$ is equivalent to the lower solution of (7) satisfying $\underline{x}_\ell \geq y_d$. Thus, $y_d \leq \underline{x}_\ell$ implies $y_d \leq \min A(y_d)$. An analogous argument shows that $y_d \geq \overline{x}_r$ implies $y_d \geq \max A(y_d)$. We have shown $y_d \notin (\underline{x}_\ell, \overline{x}_r)$ implies $y_d \notin \text{int } A(y_d)$. By contraposition, $y_d \in \text{int } A(y_d)$ implies $y_d \in (\underline{x}_\ell, \overline{x}_r)$. $\square$

A.3 Proof of Lemma 3

Lemma 2 implies Parts 1 and 5. We first prove Parts 2–4. We then prove that $\underline{x}_\ell \leq \underline{x}_r$ and $\overline{x}_\ell \leq \overline{x}_r$, which combines with Parts 1–5 to directly imply $A(y_d) \subset [\underline{x}_\ell, \overline{x}_r]$.

To prove Parts 2–4, recall that Assumption 1 implies $A(y_d) \subset (0, 1)$ for all y_d. With quadratic loss, this implies that for each $y_d \in X$ there is a ϕ such that $A(y_d) = [y_m - \phi, y_m + \phi]$ where $0 < \phi < \min\{y_m, 1 - y_m\}$ and $u(y_m, y_m \pm \phi) = \delta V_m(y_m - \phi, y_m + \phi)$. Note that $u(y_m, y_m \pm \phi) = u(0, \phi) = 1 - \phi^2$. Let

$$U_m(\phi) = \sum_{i\in N:|y_m - y_i|\leq\phi} \rho_i u(y_i, y_m) + \sum_{i\in N:|y_m - y_i|>\phi} \rho_i (1 - \phi^2).$$

Note that for an arbitrary y_d, $U_m(0) = \delta > 0$ and $U_m(\phi)$ is continuously decreasing and differentiable a.e. in ϕ.[12] Define

$$\phi(y_d) = \phi \in (0, \min\{y_m, 1 - y_m\}) \text{ such that } 1 - \phi^2 = U_m(\phi).$$

Since $\underline{x}(y_d) = y_m - \phi(y_d)$ and $\overline{x}(y_d) = y_m + \phi(y_d)$, it is sufficient to show that $\phi(y_d)$ is strictly decreasing in $|y_d - y_m|$ at a rate that approaches zero as $y_d \to y_m$ on M to prove Parts 2 and 4 of Lemma 3. Direct computations show that a.e. on M,

$$\phi'(y_d) = \frac{\delta \rho_d (y_d - y_m)}{\left(1 - \delta \sum\limits_{i \in N : |y_m - y_i| > \phi} \rho_i\right)}.$$

Note that $\phi'(y_d) < 0$ if $y_d < y_m$ and $\phi'(y_d) > 0$ if $y_m < y_d$. Since $m = \ell$ if $y_d < y_m$ and $m = r$ if $y_d > y_m$, we have that $\phi(y_d)$ is strictly decreasing in $|y_d - y_m|$ on M. Moreover, $\lim_{y_d \to^- y_\ell} \phi'(y_d) = 0$ and $\lim_{y_d \to^+ y_r} \phi'(y_d) = 0$.

To prove part 3, note that if $y_d \in (y_\ell, y_r)$, then $d = m$. Direct computations show that a.e. on (y_ℓ, y_r),

$$\phi'(y_d) = \frac{\delta \sum\limits_{i \in N : |y_m - y_i| < \phi} \rho_i (y_d - y_i)}{\left(1 - \delta \sum\limits_{i \in N : |y_m - y_i| > \phi} \rho_i\right)}.$$

Notice that

$$|\phi'(y_d)| < \frac{\delta \phi(y_d) \left(\sum\limits_{i \in N : |y_m - y_i| < \phi} \rho_i\right)}{\left(1 - \delta \sum\limits_{i \in N : |y_m - y_i| > \phi} \rho_i\right)} = \frac{\delta \phi(y_d) \left(1 - \sum\limits_{i \in N : |y_m - y_i| > \phi} \rho_i\right)}{\left(1 - \delta \sum\limits_{i \in N : |y_m - y_i| > \phi} \rho_i\right)}.$$

Because $\phi(y_d) < 1$,

$$\delta < 1 \implies \delta \left(1 - \sum_{i \in N : |y_m - y_i| > \phi} \rho_i\right) < \left(1 - \delta \sum_{i \in N : |y_m - y_i| > \phi} \rho_i\right)$$

implies $|\phi'(y_d)| \in (0, 1)$. Thus $\underline{x}'(y_d) = 1 - \phi'(y_d) > 0$ and $\overline{x}'(y_d) = 1 + \phi'(y_d) > 0$.

Next we prove that $\underline{x}_\ell \leq \underline{x}_r$ and $\overline{x}_\ell \leq \overline{x}_r$. Let

$$U_y(\zeta) = \sum_{i \in K : |y - y_i| \leq \zeta} \rho_i u(y_i, y) + \left(\rho_d + \sum_{i \in K : |y - y_i| > \zeta} \rho_i\right)(1 - \zeta^2)$$

and

$$\zeta(y) = \zeta \in (0, \min\{y, 1 - y\}) \text{ such that } 1 - \zeta^2 = U_y(\zeta).$$

Note that $\underline{x}_\ell = y_\ell - \phi(0) = y_\ell - \zeta(y_\ell)$, $\overline{x}_\ell = y_\ell + \phi(0) = y_\ell + \zeta(y_\ell)$, $\underline{x}_r = y_r - \phi(1) = y_r - \zeta(y_r)$, and $\overline{x}_r = y_r + \zeta(y_r)$. It is straightforward to check that $U_y(\zeta)$ is continuous and a.e. differentiable in $y \in [y_\ell, y_r]$ and that $\zeta(y)$ is unique for all $y \in [y_\ell, y_r]$. It is therefore sufficient to prove the result to show that $y - \zeta(y)$

---

[12] At $\phi = 0$, the acceptance set is $\{y_m\}$, so all politicians propose y_m and $V_m(y_m, y_m) = 1$. The equilibrium condition $u(y_m \pm \phi, y_m) = \delta V_m(y_m - \phi, y_m + \phi)$ then implies $U_m(\phi) = \delta V_m(y_m - \phi, y_m + \phi)$, so $U_m(0) = \delta$. Regarding differentiability, left and right derivatives of $U_m(\phi)$ with respect to y_d are unequal only at ϕ such that $|y_m - y_i| = \phi$ for some $i \in N$.

and $y + \zeta(y)$ are increasing on $[y_\ell, y_r]$. Direct computations show that wherever $\zeta(y)$ is differentiable,

$$\zeta'(y) = \frac{\delta \sum\limits_{i \in K : |y - y_i| < \zeta} \rho_i (y - y_i)}{1 - \delta \left(\rho_d + \sum\limits_{i \in K | y - y_i | > \zeta} \rho_i \right)},$$

so

$$|\zeta'(y)| < \frac{\delta \zeta(y) \left(\sum\limits_{i \in K : |y - y_i| < \zeta} \rho_i (y - y_i) \right)}{1 - \delta \left(\rho_d + \sum\limits_{i \in K | y - y_i | > \zeta} \rho_i \right)} \in (0, 1)$$

implies that $y - \zeta(y)$ and $y + \zeta(y)$ are increasing. $\square$

The following lemma characterizes when the mean policy $\mu(y_d)$ is monotone increasing, which is crucial for establishing that y_d^* is ordered.

A.4 Proof of Lemma 4.

Part 1. We first prove that $[\underline{\pi}, \overline{\pi}]$ exists. On $[y_\ell, y_r]$, both boundaries of $A(y_d)$ and y_d are strictly increasing, so $\mu(y_d)$ is strictly increasing on $[y_\ell, y_r]$. From Lemma 3, on $M_L = [\underline{x}_\ell, y_\ell]$ the median is fixed at $m = \ell$, so y_m is constant and thus $\underline{x}'(y_d) = -\phi'(y_d) = -\overline{x}'(y_d)$. Moreover, $\phi'(y_d) \to 0$ as $y_d \to^- y_\ell$. Similarly on $M_R = [y_r, \overline{x}_r]$, we have $\underline{x}'(y_d) = -\overline{x}'(y_d)$ with $\phi'(y_d) \to 0$ as $y_d \to^+ y_r$. This implies that $\mu(y_d)$ is strictly increasing on intervals $(y_\ell - \epsilon_\ell, y_\ell]$ and $[y_r, y_r + \epsilon_r)$ for $\epsilon_\ell > 0$ and $\epsilon_r > 0$. Since $A(y_d)$ is continuous, it follows that $\mu(y_d)$ is strictly increasing on $[y_\ell - \epsilon_\ell, y_r + \epsilon_r]$. Thus there are maximal $\epsilon'_\ell \in (0, y_\ell - \underline{x}_\ell]$ and $\epsilon'_r \in (0, \overline{x}_r - y_r]$ such that $\mu(y_d)$ is strictly increasing on $[y_\ell - \epsilon_\ell, y_r + \epsilon_r]$. Setting $\underline{\pi} = y_\ell - \epsilon'_\ell$ and $\overline{\pi} = y_r + \epsilon'_r$ completes the proof of Part 1.

Part 2. To prove Part 2, we show that $1 - \tilde{P}(y_\ell) > \frac{1}{2\delta}$ is a necessary condition for $\underline{\pi} > \underline{x}_\ell$. An analogous argument establishes that $\overline{\pi} < \overline{x}_r$ only if $\tilde{P}(y_\ell) > \frac{1}{2\delta}$. Since these two conditions cannot be simultaneously satisfied, it follows that $\underline{\pi} > \underline{x}_\ell$ implies $\overline{\pi} = \overline{x}_r$ and that $\overline{\pi} < \overline{x}_r$ implies $\underline{\pi} = \underline{x}_\ell$.

To show that $\underline{\pi} > \underline{x}_\ell$ implies $1 - \tilde{P}(y_\ell) > \frac{1}{2\delta}$, suppose that $\underline{\pi} > \underline{x}_\ell$. By construction, $\mu(y_d)$ is non-monotonic on $[\underline{x}_\ell, \underline{\pi}]$. Since $\mu(y_d)$ is continuous and differentiable almost everywhere, there must exist an open interval $Z \subset [\underline{x}_\ell, \underline{\pi}]$ on which

$$\begin{aligned} \mu'(y_d) &= \underline{x}'(y_d) P(\underline{x}(y_d)) + \overline{x}'(y_d)[1 - P(\overline{x}(y_d))] + \rho_d \\ &= \delta \rho_d (y_\ell - y_d) \left(\frac{P(\underline{x}(y_d)) - [1 - P(\overline{x}(y_d))]}{1 - \delta[P(\underline{x}(y_d)) + 1 - P(\overline{x}(y_d))]} \right) + \rho_d < 0 \\ \implies\ 1 &< \delta (y_\ell - y_d) \left(\frac{[1 - P(\overline{x}(y_d))] - P(\underline{x}(y_d))}{1 - \delta[P(\underline{x}(y_d)) + 1 - P(\overline{x}(y_d))]} \right). \end{aligned}$$

Since $0 < y_\ell - y_d < 1$, it must be that $P(\underline{x}(y_d)) < [1 - P(\overline{x}(y_d))]$ and

$$1 < \delta \left(\frac{[1 - P(\overline{x}(y_d))] - P(\underline{x}(y_d))}{1 - \delta[P(\underline{x}(y_d)) + 1 - P(\overline{x}(y_d))]} \right),$$

which implies

$$\frac{1}{2\delta} < [1 - P(\overline{x}(y_d))].$$

Because $y_\ell < \overline{x}(y_d)$, the last inequality can be satisfied for some $y_d \in M_L$ only if $\frac{1}{2\delta} < 1 - \tilde{P}(y_\ell)$.

Part 3. We prove the result for an arbitrary interval $Z \subset M_L$ on which $\mu(y_d)$ is decreasing. The proof for $Z \subset M_R$ is analogous. We first show that $y_\ell < \mu(y_d)$. We know from Part 2 of Lemma 4 that $1 - P(\overline{x}(y_d)) > \frac{1}{2\delta} > \frac{1}{2}$ at any $y_d \in M_L$ such that $\mu(y_d)$ is non-increasing. Thus for $y_d \in Z$,

$$\begin{aligned}
\mu(y_d) &= \underline{x}(y_d)P(\underline{x}(y_d)) + \overline{x}(y_d)[1 - P(\overline{x}(y_d))] + \sum_{i \in N : y_i \in (\underline{x}(y_d), \overline{x}(y_d)]} \rho_i y_i \\
&> \underline{x}(y_d)[1 - (1 - P(\overline{x}(y_d)))] + \overline{x}(y_d)[1 - P(\overline{x}(y_d))] \\
&= [y_\ell - \phi(y_d)]P(\overline{x}(y_d)) + [y_\ell + \phi(y_d)][1 - P(\overline{x}(y_d))] \\
&= y_\ell + \phi(y_d)[1 - 2P(\overline{x}(y_d))] \\
&> y_\ell.
\end{aligned}$$

Next we show that $\sigma^2(y_d)$ is strictly decreasing in y_d on Z. Because $\mu(y_d)$ is continuous and a.e. differentiable, $\sigma^2(y_d)$ must be too. Since $y_d \in M_L$,

$$u(\overline{x}(y_d), y_\ell) = \delta V_\ell(\underline{x}(y_d), \overline{x}(y_d))$$

$$\implies 1 - (\overline{x}(y_d) - y_\ell)^2 = \delta[1 - (\mu(y_d) - y_\ell)^2 - \sigma^2(y_d)]$$

$$\implies \delta\sigma^2(y_d) = \delta[1 - (y_\ell - \mu(y_d))^2] - [1 - (\overline{x}(y_d) - y_\ell)^2].$$

Thus almost everywhere on Z,

$$\frac{\partial \delta\sigma^2(y_d)}{\partial y_d} = -2\delta[\mu(y_d) - y_\ell]\mu'(y_d) + 2[\overline{x}(y_d) - y_\ell]\overline{x}'(y_d) < 0$$

if

$$\overline{x}'(y_d)[\overline{x}(y_d) - y_\ell] - \delta\mu'(y_d)[\mu(y_d) - y_\ell] < 0. \tag{9}$$

We already showed $\mu'(y_d) \leq 0$ implies $y_\ell < \mu(y_d) < \overline{x}(y_d)$. Therefore

$$\overline{x}'(y_d) - \delta\mu'(y_d) < 0 \tag{10}$$

implies (9). Substituting

$$\mu'(y_d) = \overline{x}'(y_d)[1 - P(\overline{x}(y_d))] + \underline{x}'(y_d)P(\underline{x}(y_d)) + \rho_d = \overline{x}'(y_d)[1 - P(\overline{x}(y_d)) - P(\underline{x}(y_d))] + \rho_d$$

into (10) yields

$$\overline{x}'(y_d) - \delta\mu'(y_d) = \overline{x}'(y_d)[1 - \delta(1 - P(\overline{x}(y_d)) - P(\underline{x}(y_d)))] - \delta\rho_d < 0.$$

Because $\overline{x}'(y_d) < 0$ by Lemma 3 and $[1 - P(\overline{x}(y_d)) - P(\underline{x}(y_d))] \in (0, 1)$, its follows that (10) is satisfied. Thus $\sigma^2(y_d)$ is almost everywhere strictly decreasing on Z. The continuity of $\sigma^2(y_d)$ then implies that $\sigma^2(y_d)$ is strictly decreasing on Z. □

B Optimal Representatives

We establish general properties of $y_d^*(y_p)$ in Lemma A1. We use Lemmas A2-A4 to prove that $y_d^*(y_d)$ is ordered on X and single-valued almost everywhere on (E_L, E_R).

B.1 General Properties of y_d^*

Lemma A1.

1. $[0,\underline{x}_\ell] \cap y_d^*(y_p) \neq \emptyset \iff [0,\underline{x}_\ell] \subseteq y_d^*(y_p)$ *and* $[\overline{x}_r, 1] \cap y_d^*(y_p) \neq \emptyset \iff [\overline{x}_r, 1] \subseteq y_d^*(y_p)$*;*
2. $y_p \geq y_\ell$ *implies* $y_d^*(y_p) \cap [\underline{x}_\ell, y_\ell) = \emptyset$*, and* $y_p \leq y_r$ *implies* $y_d^*(y_p) \cap (y_r, \overline{x}_r] = \emptyset$*;*
3. $y_p < y_\ell$ *implies* $y_\ell \notin y_d^*(y_p)$*, and* $y_p > y_r$ *implies* $y_r \notin y_d^*(y_p)$*;*
4. $y_p \in (\underline{x}_\ell, y_\ell)$ *implies* $[\underline{x}_\ell, y_p] \cap y_d^*(y_p) = \emptyset$*, and* $y_p \in (y_r, \overline{x}_r)$ *implies* $[y_p, \overline{x}_r] \cap y_d^*(y_p) = \emptyset$*; and*
5. $y_p \in [0, \underline{x}(y_\ell)] \cup [\overline{x}(y_r), 1]$ *implies* $[y_\ell, y_r] \cap y_d^*(y_p) = \emptyset$*.*
6. *There exist* $e_L < \underline{x}_\ell$ *and* $e_R > \overline{x}_r$ *such that* $y_d^*(y_p) \subset (\underline{x}_\ell, \overline{x}_r)$ *if* $y_d \in (e_L, e_R)$*.*

Proof.
The principal solves $\max_{y_d \in [0,1]} U(y_d, y_p)$ where $U : [0,1]^2 \to \mathbb{R}$ is:

$$\begin{aligned} U(y_d, y_p) &\equiv P(\underline{x}(y_d))u(\underline{x}(y_d), y_p) + [1 - P(\overline{x}(y_d))]u(\overline{x}(y_d), y_p) + \sum_{i \in N : y_i \in (\underline{x}(y_d), \overline{x}(y_d)]} \rho_i u(y_i, y_p) \\ &= 1 - (y_p - \mu(y_d))^2 - \sigma^2(y_d). \end{aligned}$$

For each part 1–5, we prove one side since the other side is analogous.

1. For $y_d \leq \underline{x}_\ell$, Lemma 3 implies $A(y_d) = [\underline{x}_\ell, \overline{x}_\ell]$, so $U(\underline{x}_\ell; y_p) = U(y_d; y_p)$.
2. Consider $y_p \leq y_r$ and suppose there exists $y_d' \in (y_r, \overline{x}_r]$ such that $y_d' \in y_d^*(y_p)$. We establish a contradiction by showing there must exist $y_d'' < y_r$ such that $U(y''; y_p) > U(y'; y_p)$. First, Lemma 2 implies $y_p < y_d' < \overline{x}(y_d')$.

   Because $\underline{x}(y_d)$ is strictly decreasing and $\overline{x}(y_d)$ strictly increasing on $[y_r, \overline{x}_r]$, we know $y_d' \in y_d^*(y_p)$ requires $y_p \leq \underline{x}(y_d')$ (otherwise, $U(y_d - \varepsilon; y_p) > U(y_d; y_p)$ for some $\varepsilon > 0$). Next, by Lemma 3, we know that (i) $\underline{x}(y_d)$ is continuous and strictly increasing on $[\underline{x}_\ell, y_r]$, (ii) $\underline{x}_\ell < \underline{x}_r$, and (iii) $y_d \leq \underline{x}(y_d)$ if and only if $y_d \leq \underline{x}_\ell$. Thus, a $y_d'' \in (y_p, y_r)$ exists such that $\underline{x}(y_d'') = \underline{x}(y_d')$. Since $\overline{x}(y_d) = 2y_m - \underline{x}(y_d)$ and $y_p < y_d'' < y_r < y_d'$, it follows that $y_p < \overline{x}(y_d'') < 2y_r - \underline{x}(y') = \overline{x}(y_d')$. We have shown that $|y_p - y_d''| < |y_p - y_d'|$, $|y_p - \overline{x}(y_d'')| < |y_p - \overline{x}(y_d')|$, and $|y_p - \underline{x}(y_d'')| = |y_p - \underline{x}(y_d')|$, which implies $U(y_d', y_p) < U(y_d'', y_p)$. But then $y_p' \notin y_d^*(y_p)$, a contradiction.
3. Lemma 3 implies that for $y_p < y_\ell$, $U(y_d; y_p)$ is continuously differentiable on $(y_\ell - \varepsilon, y_\ell)$ with

   $$\lim_{y_d \to y_\ell^-} \frac{\partial U(y_d; y_p)}{\partial y_d} = \rho_d \cdot \frac{\partial u(y_d, y_p)}{y_d}\Big|_{y_d = y_\ell^-} < 0.$$

   Then, continuity of $U(y_d; y_p)$ implies $y_\ell \notin y_d^*(y_p)$.

4. Consider $y_p \in (\underline{x}_\ell, y_\ell)$. Then, $y_p \in \text{int}A(y_d)$ for all $y_d \leq y_p$. By Lemma 3, $\underline{x}(y_d)$ is continuous and strictly increasing on $[\underline{x}_\ell, y_\ell]$, while $\overline{x}(y_d)$ is continuous and strictly decreasing. Thus, there exists $\varepsilon > 0$ such that $P(\underline{x}(y_d))\, u(\underline{x}(y_d), y_p) + [1 - P(\overline{x}(y_d))]\, u(\overline{x}(y_d), y_p)$ is strictly increasing on $[\underline{x}_\ell, y_p + \varepsilon]$. By assumption, $\frac{\partial u(y_d;y_p)}{\partial y_d} > 0$ if $y_d < y_p$ and $\frac{\partial u(y_d;y_p)}{\partial y_d}\big|_{y_d = y_p} = 0$. By continuity of $\frac{\partial u(y_d;y_p)}{\partial y_d}$ there exists $y_d' \in (y_p, y_p + \varepsilon)$ such that $U(y_d; y_p)$ is strictly increasing over $y_d \in [\underline{x}_\ell, y_d')$.

5. Lemma 3 establishes that $\underline{x}(y_d)$ and $\overline{x}(y_d)$ are strictly increasing on $[y_\ell, y_r]$. Therefore $U(y_d; y_p)$ is strictly decreasing on $[y_\ell, y_r]$ if $y_p \leq \underline{x}(y_\ell)$. Thus $\text{argmax}_{y_d \in [y_\ell, y_r]} U(y_d; y_p) = y_\ell$ for all $y_p \leq \underline{x}(y_r)$. But by part 3, $y_\ell \in y_d^*(y_p)$ only if $y_p \geq y_\ell$.

6. Parts 1–5 imply that $y_d^*(y_p) \subset (\underline{x}_\ell, \overline{x}_r)$ for all $y_p \in [\underline{x}_\ell, \overline{x}_r]$. Part 6 then follows from upper hemicontinuity of y_d^*. □

B.2 Ordering of y_d^*

Lemma A2. *Let $Y \subseteq X$ denote an arbitrary subset of X and define $\tilde{y}_d : X \times 2^X \to 2^X$ as*

$$\tilde{y}_d(y_p; Y) \equiv \underset{y_d \in Y}{\text{argmax}}\, U(y_d, y_p).$$

If $\mu(y_d)$ is weakly increasing on Y, then $\tilde{y}_d(y_p; Y)$ is increasing in the strong set order sense. If $\mu(y_d)$ is strictly increasing on Y, then every selection from $\tilde{y}_d(y_p; Y)$ is weakly increasing.

Proof. Consider an arbitrary subset $Y \subseteq X$ and arbitrary $y_d, y_d' \in Y$. Since

$$U(y_d', y_p) - U(y_d, y_p) = 1 - (\mu(y_d') - y_p)^2 - \sigma^2(y_d') - [1 - (\mu(y_d) - y_p)^2 - \sigma^2(y_d))]$$

$$\implies \frac{\partial [U(y_d', y_p) - U(y_d, y_p)]}{\partial y_p} \propto \mu(y_d') - \mu(y_d),$$

we have the following:

1. If $y_d < y_d' \implies \mu(y_d) \leq \mu(y_d')$, then $U(y_d, y_p)$ satisfies the single-crossing property.

2. If $y_d < y_d' \implies \mu(y_d) < \mu(y_d')$, then $U(y_d, y_p)$ satisfies the strict single-crossing property.

The result then follows directly from Theorem 4 in Milgrom and Shannon (1994). □

Lemma A3. *$\mu(y_d)$ is strictly increasing on $\mathcal{M}_L \cup [y_\ell, y_r] \cup \mathcal{M}_R$ and weakly increasing on D where*

$$\mathcal{M}_L \equiv \{y_d \in M_L : \mu(y_d) < \mu(y_d') \textit{ for all } y_d' \in M_L \textit{ such that } y_d < y_d'\},$$

$$\mathcal{M}_R \equiv \{y_d \in M_R : \mu(y_d') < \mu(y_d) \textit{ for all } y_d' \in M_R \textit{ such that } y_d' < y_d\},$$

and

$$D \equiv \begin{cases} \mathcal{M}_L \cup [y_\ell, y_r] \cup \mathcal{M}_R \cup [\overline{x}_r, 1] & \textit{if} \quad \underline{x}_\ell \notin \mathcal{M}_L \\ [0, \underline{x}_\ell] \cup \mathcal{M}_L \cup [y_\ell, y_r] \cup \mathcal{M}_R & \textit{if} \quad \overline{x}_r \notin \mathcal{M}_R \\ [0, \underline{x}_\ell] \cup \mathcal{M}_L \cup [y_\ell, y_r] \cup \mathcal{M}_R \cup [\overline{x}_r, 1] & \textit{otherwise} \end{cases}$$

Proof. We first show that $\mu(y_d)$ strictly increases on $\mathcal{M}_L \cup [y_\ell, y_r] \cup \mathcal{M}_R$. Both boundaries of the acceptance set strictly increase on $[y_\ell, y_r]$ which implies that $\mu(y_d)$ is strictly increasing on $[y_\ell, y_r]$. Lemma 4 implies that $y_\ell \in \mathcal{M}_L$ and $y_r \in \mathcal{M}_R$. Since $\mu(y_d)$ is strictly increasing on $\mathcal{M}_L$ and $\mathcal{M}_R$, it follows that $\mu(y_d)$ strictly increases on $\mathcal{M}_L \cup [y_\ell, y_r] \cup \mathcal{M}_R$.

We now show that $\mu(y_d)$ is weakly increasing on D. Since $A(y_d)$ is constant on $[0, \underline{x}_\ell]$ and $[\overline{x}_r, 1]$, so is $\mu(y_d)$. Thus if $\underline{x}_\ell \in \mathcal{M}_L$ and $\overline{x}_r \in \mathcal{M}_R$, then (i) $\mu(y_d) = \mu(\underline{x}_\ell) < \mu(y_d')$ for all $y_d \leq \underline{x}_r$ and $y_d' \in D$ such that $y_d' > \underline{x}_\ell$ and (ii) $\mu(y_d') < \mu(\overline{x}_r) = \mu(y_d)$ for all $y_d \geq \overline{x}_r$ and $y_d' \in D$ such that $y_d' < \overline{x}_r$. If $\underline{x}_\ell \notin \mathcal{M}_L$, then $\overline{x}_r \in \mathcal{M}_R$ by Lemma 4 so $\mu(y_d)$ is strictly increasing on $D \setminus (\overline{x}_r, 1]$ and $\mu(y_d) = \mu(\overline{x}_r)$ for all $y_d > \overline{x}_r$. Analogously, if $\overline{x}_r \notin \mathcal{M}_R$, then $\underline{x}_\ell \in \mathcal{M}_L$ so $\mu(y_d)$ strictly increases on $D \setminus [0, \underline{x}_\ell)$ and $\mu(y_d) = \mu(\underline{x}_r)$ for all $y_d < \underline{x}_\ell$. $\square$

Lemma A4. *The image of y_d^* is a subset of D*

Proof. Let $\tilde{Y} = \{y_d^*(y_p) | y_p \in X\}$ denote the image of y_d^* and define $\mathcal{M} \equiv \mathcal{M}_L \cup \mathcal{M}_R$. Lemma A1 implies $\underline{x}_\ell \in \tilde{Y} \iff [0, \underline{x}_\ell] \subset \tilde{Y}$ and $\overline{x}_r \in \tilde{Y} \iff [\overline{x}_r, 1] \subset \tilde{Y}$. By definition, $\underline{x}_\ell \in \mathcal{M} \iff [0, \underline{x}_\ell] \subset D$ and $\overline{x}_r \in \mathcal{M} \iff [\overline{x}_r, 1] \subset D$. We can therefore prove Lemma A4 by showing that $\tilde{Y} \cap \{M \setminus \mathcal{M}\} = \emptyset$.

By definition, $y_d \in \{M \setminus \mathcal{M}\}$ if and only if one of the following holds:

1. If $y_d \in M_L$, then a $y_d' \in M_L$ exists such that $y_d < y_d'$ and either
   (a) $\mu(y_d') = \mu(y_d)$, or
   (b) $\mu(y_d') < \mu(y_d)$.
2. If $y_d \in M_R$, then a $y_d'' \in M_R$ exists such that $y_d'' < y_d$ and either
   (a) $\mu(y_d) = \mu(y_d'')$, or
   (b) $\mu(y_d) < \mu(y_d'')$.

We show $y_d \notin \tilde{Y}$ for all of these cases. We proceed in two steps.

Step 1. We first show that $y_d \notin \tilde{Y}$ if y_d satisfies condition 1(a) or 2(a). To do so, we prove the following:

Claim: *Suppose $y_d < y_d'$ and $\mu(y_d) = \mu(y_d')$. Then for all y_p, $U(y_d, y_p) < U(y_d', y_p)$ if $y_d, y_d' \in M_L$ and $U(y_d', y_p) < U(y_d, y_p)$ if $y_d, y_d' \in M_R$.*

Proof: We prove the claim for $y_d', y_d \in M_L$. An analogous argument establishes the result for $y_d', y_d \in M_R$. For this case, we know $m = \ell$'s continuation value, $V_\ell = 1 - (\mu(y_d) - y_\ell)^2 - \sigma^2(y_d)$, is strictly increasing in y_d on M_L since $\underline{x}(y_d)$ is strictly increasing and $\overline{x}(y_d)$ strictly decreasing. Thus $\mu(y_d) = \mu(y_d')$ implies $\sigma^2(y_d') < \sigma^2(y_d)$. Therefore $U(y_d', y_p) - U(y_d, y_p) = \sigma^2(y_d) - \sigma^2(y_d') > 0$ for all y_p. $\blacksquare$

Step 2. We now show that $y_d \notin y_d^*(X)$ if y_d satisfies condition 1(b) or 2(b). Part 1 of Lemma 4 implies that $M \setminus \mathcal{M} \subseteq [\underline{x}_\ell, y_\ell) \cup (y_r, \overline{x}_r]$. Lemma A1 establishes that $y_d^*(y_p) \cap [\underline{x}_\ell, y_\ell) = \emptyset$ if $y_p \geq y_\ell$ and $y_d^*(y_p) \cap (y_r, \overline{x}_r] = \emptyset$ if $y_p \leq y_r$. It is therefore sufficient to show that $y_d \notin y_d^*(y_p)$ for any $y_p < y_\ell$ if y_d satisfies 1(b) and that $y_d \notin y_d^*(y_p)$ for any $y_p > y_r$ if y_d satisfies 2(b). To show this, we prove the following:

Claim: *Consider an arbitrary interval $Z = [\underline{z}, \overline{z}] \subset M$ on which $\mu(y_d)$ is decreasing.*

1. *If $Z \subset M_L$, then* $\operatorname*{argmax}_{y_d \in Z} U(y_d, y_p) = \overline{z}$ *for all $y_p < y_\ell$.*

2. *If* $Z \subset M_R$*, then* $\underset{y_d \in Z}{\operatorname{argmax}} U(y_d, y_p) = \underline{z}$ *for all* $y_p > y_r$*.*

Proof: We prove the first part; the second part is analogous. Consider an arbitrary interval $Z = [\underline{z}, \overline{z}] \subseteq M_L$ such that $\mu(y_d)$ is non-increasing on Z. Since $\mu(y_d)$ is strictly increasing on $[\underline{\pi}, \overline{\pi}]$, it must be that $Z \subset [\underline{x}_\ell, \underline{\pi}]$. Part 3 of Lemma 4 therefore implies that for all $y_d \in Z$, (i) $y_\ell < \mu(y_d)$ and (ii) $\sigma^2(y_d)$ is strictly decreasing. Thus $U(y_d, y_p) = 1 - (\mu(y_d) - y_p)^2 - \sigma^2(y_d)$ is strictly increasing on Z for all $y_p < y_\ell$. ■ □

Lemma A5. *The correspondence* y_d^* *is increasing in* $y_p \in X$ *in the strong set order sense and* $y_d^*|_{(E_L, E_R)}$ *is a singleton almost everywhere.*

Proof. Lemma A4 implies that $y_d^*(y_p) = \tilde{y}_d(y_p; D)$. Since $\mu(y_d)$ is increasing on D by Remark A3, Lemma A2 implies that $\tilde{y}_d(y_p; D)$ is increasing in the strong set order sense, so $y_d^*(y_p)$ is too. Lemmas A1 and A4 imply $y_d^*|_{(E_L, E_R)}(y_p) = \tilde{y}_d|_{(E_L, E_R)}(y_p; \mathcal{M} \cup [y_\ell, y_r])$. Since $\mu(y_d)$ is strictly increasing on $\mathcal{M} \cup [y_\ell, y_r]$ by Remark A3, Lemma A2 implies that every selection from $\tilde{y}_d|_{(E_L, E_R)}(y_p; \mathcal{M} \cup [y_\ell, y_r])$ is increasing and therefore the same must be true of $y_d^*|_{(E_L, E_R)}(y_p)$. Since y_d^* is upper hemicontinuous and increasing in the strong set order sense, it follows that $y_d^*|_{(E_L, E_R)}(y_p)$ is a singleton almost everywhere (Kenderov 1976). □

B.3 Proof of Proposition 1

Follows immediately from Lemmas A1–A5. □

C Polarized Legislature

First, in Lemma A6 we refine our characterization of $\underline{x}(y_d)$ and $\overline{x}(y_d)$ under Assumptions 1 and 2. In Lemmas A7 and A8, we characterize the principal's locally optimal representative within the set of centrist representatives and aligned moderate representatives, respectively. We use these to prove Propositions A.1–A.3.

Lemma A6. *Suppose Assumptions 1 and 2. Then:*

$$\underline{x}_\ell = y_\ell - \sqrt{\frac{1 - \delta + \delta\rho_r(y_\ell - y_r)^2}{1 - \delta(\rho_E + \rho_d)}}, \text{ and} \tag{11}$$

$$\overline{x}_r = y_r + \sqrt{\frac{1 - \delta + \delta\rho_\ell(y_\ell - y_r)^2}{1 - \delta(\rho_E + \rho_d)}}. \tag{12}$$

Furthermore, (i) $\underline{x}$ *and* $\overline{x}$ *are* $\mathcal{C}^2$ *on* $(\underline{x}_\ell, y_\ell) \cup (y_\ell, y_r) \cup (y_r, \overline{x}_r)$*; (ii)* $\underline{x}(y_d)$ *is strictly concave and* $\overline{x}(y_d)$ *strictly convex on each of those intervals; and (iii)* $\frac{\underline{x}'(y_d)}{\overline{x}'(y_d)}$ *is strictly decreasing over* $y_d \in (y_\ell, y_r)$*, with* $\frac{\underline{x}'(y_d)}{\overline{x}'(y_d)} = 1$ *if and only if* $y_d = \frac{\rho_\ell y_\ell + \rho_r y_r}{\rho_\ell + \rho_r} \in (y_\ell, y_r)$*.*

Proof. Direct computations yield $\underline{x}(y_d) = y_m - \phi(y_d)$ and $\overline{x}(y_d) = y_m + \phi(y_d)$ for each $y_d \in (\underline{x}_\ell, \overline{x}_r)$, where:

$$\phi(y_d) = \sqrt{\frac{1 - \delta + \delta\rho_\ell(y_\ell - y_m)^2 + \delta\rho_r(y_r - y_m)^2 + \delta\rho_d(y_d - y_m)^2}{1 - \delta\rho_E}}.$$

First, (11) follows from solving $y_d = \underline{x}(y_d)$ for $y_d < y_\ell$ and similarly (12) follows from solving $y_d = \overline{x}(y_d)$ for $y_d > y_r$. Next, $[\phi(y_d)]^2$ is a quadratic polynomial with a positive leading coefficient on each of the intervals $(\underline{x}_\ell, y_\ell)$, (y_ℓ, y_r), and $(y_r, \overline{x}_r)$, so it is strictly convex on each interval. Thus, on each interval $\underline{x}$ is strictly concave and $\overline{x}$ is strictly convex. Finally, direct computations yield that $\frac{\underline{x}'(y_d)}{\overline{x}'(y_d)}$ is strictly decreasing over $y_d \in (y_\ell, y_r)$, with $\frac{\underline{x}'(y_d)}{\overline{x}'(y_d)} = 1$ if and only if $y_d = \frac{\rho_\ell y_\ell + \rho_r y_r}{\rho_\ell + \rho_r}$. □

Lemma A7. *Under Assumptions 1 and 2, the mapping* $\hat{y}_d(y_p) \equiv \underset{y_d \in [y_\ell, y_r]}{\operatorname{argmax}} U(y_d; y_p)$ *is equivalent to a function* $\hat{y}_d : X \to [y_\ell, y_r]$ *that is continuous and weakly increasing. Furthermore, (i)* $\hat{y}_d|_{[y_\ell, y_r]}$ *has a unique fixed point* y^**, (ii)* $y_p < y^*$ *implies* $\hat{y}_d(y_p) \in (y_p, y^*]$*, and (iii)* $y_p > y^*$ *implies* $\hat{y}_d(y_p) \in [y^*, y_p)$.

Proof. By Lemma A6, for all $y_d \in (y_\ell, y_r)$ we have: $\underline{x}''(y_d) < 0 < \underline{x}'(y_d)$ and $0 < \min\{\overline{x}'(y_d), \overline{x}''(y_d)\}$. Thus for an arbitrary $y_p \in X$ and $y_d \in (y_\ell, y_r)$, (i) $y_p \geq \overline{x}(y_d)$ implies $\frac{\partial U(y_d; y_p)}{\partial y_d} > 0$, (ii) $y_p \in (\underline{x}(y_d), \overline{x}(y_d))$ implies $\frac{\partial^2 U(y_d; y_p)}{\partial y_d^2} < 0$, and (iii) $y_p \leq \underline{x}(y_d)$ implies $\frac{\partial U(y_d; y_p)}{\partial y_d} < 0$, so $U(y_d; y_p)$ is strictly quasi-concave over $y_d \in [y_\ell, y_r]$ for all $y_p \in X$. Additionally, $\mu(y_d)$ is strictly increasing on $[y_\ell, y_r]$, so $U(y_d; y_p)$ satisfies the strict single-crossing condition on $[y_\ell, y_r] \times X$. Thus, $\hat{y}_d(y_p)$ is single-valued, continuous, and increasing. It follows that $\hat{y}_d|_{[y_\ell, y_r]}$ has a fixed point. To show it is unique, first note that: (i) $\hat{y}_d(y_\ell) = y_\ell$ if and only if $\frac{\partial U(y_d; y_\ell)}{\partial y_d}\Big|_{y_d = y_\ell^+} \leq 0$, (ii) $\hat{y}_d(y_r) = y_r$ if and only if $\frac{\partial U(y_d; y_r)}{\partial y_d}\Big|_{y_d = y_r^-} \geq 0$, and (iii) and $\hat{y}_d(y_p) = y_p \in (y_\ell, y_r)$ if and only if $\frac{\partial U(y_d; y_p)}{\partial y_d}\Big|_{y_d = y_p} = 0$. In the main text we define

$$\lambda(y_p) := \frac{\partial U(y_d, y_p)}{\partial y_d}\Big|_{y_d = y_p} = \rho_L \frac{\partial \underline{x}(y_d)}{\partial y_d}\Big|_{y_d = y_p} - \rho_R \frac{\partial \overline{x}(y_d)}{\partial y_d}\Big|_{y_d = y_p}.$$

An interior fixed point exists if and only if $\lambda(y_p) = 0$ for some $y_p \in (y_\ell, y_r)$. Strict concavity of $\underline{x}(y_d)$ and strict convexity of $\overline{x}(y_d)$ imply $\lambda'(y_p) < 0$. Therefore $\lambda(y_p) = 0$ at most once, which implies that y^* is unique. Furthermore, $y_p^* \in \{y_\ell, y_r\}$ if λ does not change sign, and otherwise $y^* \in (y_\ell, y_r)$. Finally, since $\hat{y}_d(y_p)$ weakly increasing and $\lambda'(y_p) < 0$, we know that (i) $y_p < y_p^*$ implies $\hat{y}_d(y_p) \in (y_p, y^*]$, and (ii) $y_p > y^*$ implies $\hat{y}_d(y_p) \in [y^*, y_p)$. □

Lemma A8. *Define* $\tilde{y}_d : [0, y_\ell] \to [\underline{x}_\ell, y_\ell]$ *as* $\tilde{y}_d(y_p) \equiv \operatorname{argmax}_{y_d \in [\underline{x}_\ell, y_\ell]} U(y_d, y_p)$*. Under Assumptions 1-2,* $\tilde{y}_d$ *is single-valued, continuous, increasing, and satisfies* $\tilde{y}_d(y_p) \in [\underline{\pi}, y_\ell]$ *for all* $y_p \leq y_\ell$*. Furthermore,* $\tilde{y}_d(y_\ell) = y_\ell$ *and otherwise* $\tilde{y}_d(y_p) \in (y_p, y_\ell)$*. A unique* $E_L < \underline{x}_\ell$ *exists such that (i)* $\tilde{y}_d(y_p) = \underline{x}_\ell$ *if* $y_p < E_L$*, (ii)* $\tilde{y}_d(y_p) > \underline{x}_\ell$ *if* $y_p > E_L$*, and (iii)* $\tilde{y}_d(y_p)$ *is strictly increasing on* (E_L, y_ℓ)*. For* $y_p \geq y_r$*, analogous properties hold for* $\operatorname{argmax}_{y_d \in [y_r, \overline{x}_r]} U(y_d, y_p)$.

Proof. Fix $y_p \leq y_\ell$. To begin, we show that $\tilde{y}_d(y_p) \subset [\underline{\pi}, y_\ell]$. The result is trivial if $\underline{\pi} = \underline{x}_\ell$. If $\underline{\pi} > \underline{x}_\ell$, then an interval $Z \subseteq [\underline{x}_\ell, \underline{\pi}]$ exists where $\mu(y_d)$ is decreasing in y_d. Moreover, $\rho_R > 1/2\delta$ by Lemma 4. Since $A(y_d)$ is continuously differentiable on M_L under Assumption 2 at rates of change that decrease continuously in magnitude by Lemma 3, it follows that $\mu(y_d)$ is strictly decreasing on $[\underline{x}_\ell, \underline{\pi})$. Thus Lemma 4 implies $U(y_d, y_p)$ is strictly increasing in y_d on $[\underline{x}_\ell, \underline{\pi})$ for all $y_p \leq y_\ell$ so $\tilde{y}_d(y_p) \subset [\underline{\pi}, y_\ell]$.

Since $\mu(y_d)$ strictly increases on $[\underline{\pi}, y_\ell]$, it follows that $U(y_d, y_p)$ satisfies strict single crossing. Thus every selection from $\tilde{y}_d(y_p)$ must be increasing, in addition to $\tilde{y}_d(y_p)$ being non-empty, upper hemicontinuous, and compact-valued by Berge's maximum theorem. Furthermore, Lemma A1 implies that $\tilde{y}_d(y_\ell) = y_\ell$ and otherwise $\tilde{y}_d(y_p) \subset (y_p, y_\ell)$. Thus, there is a unique $E_L \equiv \inf\{y_p < \underline{x}_\ell : \tilde{y}_d(y_p) \subset (\underline{x}_\ell, y_\ell)\}$. For all $y_p \in (E_L, y_\ell)$, any $y_d \in \tilde{y}_d(y_p)$ must satisfy

$$\frac{\partial U(y_d; y_p)}{\partial y_d} \propto -\rho_d(y_d - y_p) - \underline{x}'(y_d)[\rho_L(\underline{x}(y_d) - y_p) - \rho_R(\overline{x}(y_d) - y_p)] = 0 \tag{13}$$

and

$$\frac{\partial^2 U(y_d; y_p)}{\partial y_d^2} \propto -\rho_d - \underline{x}''(y_d)[\rho_L(\underline{x}(y_d) - y_p) - \rho_R(\overline{x}(y_d) - y_p)] - [\underline{x}'(y_d)]^2 \rho_E < 0. \tag{14}$$

Since $y_p < \tilde{y}_d(y_p)$ and $\underline{x}'(y_d) > 0$, (13) holds only if $\rho_L(\underline{x}(y_d) - y_p) - \rho_R(\overline{x}(y_d) - y_p) < 0$.

To show $\tilde{y}_d(y_p)$ is single-valued for all $y_p \in (E_L, y_\ell)$, suppose not and let $y_d, y_d' \in \tilde{y}_d(y_p)$ where $y_d < y_d'$. Then, there must be a $y \in (y_d, y_d')$ satisfying $\rho_L(\underline{x}(y) - y_p) - \rho_R(\overline{x}(y) - y_p) = -\left(\frac{\rho_d + \rho_E[\underline{x}'(y)]^2}{\underline{x}''(y)}\right) > 0$. And since $\frac{\partial}{\partial y}[\rho_L(\underline{x}(y) - y_p) - \rho_R(\overline{x}(y) - y_p)] = \underline{x}'(y)\rho_E > 0$, we must also have $\rho_L(\underline{x}(y_d') - y_p) - \rho_R(\overline{x}(y_d') - y_p) > 0$. But then (13) fails at y_d' which implies that $y_d' \notin \tilde{y}_d(y_p)$, contradicting our assumption that $y_d' \in \tilde{y}_d(y_p)$. Consequently, $\tilde{y}_d(y_p)$ is single-valued. Finally, applying the implicit function theorem to $\tilde{y}_d(y_p)$ shows that it is strictly increasing on $[E_L, y_p]$. Analogous arguments establish the result for $y_p \geq y_r$. □

Proposition A.1 (Unique Locus of Attraction)**.** *Suppose Assumptions 1-2. Over $y_p \in (E_L, E_R)$, the optimal representative correspondence y_d^* has a unique fixed point, y^*. Moreover, (i) $\lambda(y_\ell) \leq 0$ implies $y^* = y_\ell$; (ii) $\lambda(y_r) \geq 0$ implies $y_r = y^*$; and (iii) otherwise, $y^* \in (y_\ell, y_r)$.*

Proof. First, Lemma A1 implies that any fixed point of $y_d^*|_{(E_L,E_R)}$ must be in $[y_\ell, y_r]$, where $y_d^*(y_p) = \hat{y}_d(y_p)$. Second, by Lemma A7: $y_d^*|_{[y_\ell,y_r]}$ has a unique fixed point; $y_p < y^*$ implies $\hat{y}_d(y_p) \in (y_p, y^*]$; and $y_p > y^*$ implies $\hat{y}_d(y_p) \in [y^*, y_p)$. Third, by Proposition 1: $y_p \in (E_L, \underline{y}_p)$ implies $y_d^*(y_p) \in (y_p, y_\ell)$; $y_p \in (\underline{y}_p, \overline{y}_p)$ implies $y_d^*(y_p) = \hat{y}_d(y_p)$; and $y_p \in (\overline{y}_p, E_R)$ implies $y_d^*(y_p) \in (y_r, y_p)$. Thus, y^* is the unique fixed point of $y_d^*|_{(E_L,E_R)}$. □

Proposition A.2 (Dead Zone Representatives)**.** *Suppose Assumptions 1-2. In equilibrium,*

1. *$\lambda(y_\ell) \leq 0$ implies $\underline{y}_p = y_\ell < y_r < \overline{y}_p$, so $y_r \in \Delta$ but y_ℓ is not;*
2. *$\lambda(y_r) \geq 0$ implies $\underline{y}_p < y_\ell < y_r = \overline{y}_p$, so $y_\ell \in \Delta$ but y_r is not; and*
3. *otherwise, $\underline{y}_p < y_\ell < y_r < \overline{y}_p$, so $\{y_\ell, y_r\} \subset \Delta$.*

Proof. By Lemma A1, (i) $y_p < y_\ell$ implies $y_\ell \notin y_d^*(y_p)$ and (ii) $y_p > y_r$ implies $y_r \notin y_d^*(y_p)$. Then, since $y_d^*(y_p)$ is increasing, (i) $\underline{y}_p = y_\ell$ if and only if $y^* = y_\ell$ and (ii) $\overline{y}_p = y_r$ if and only if $y^* = y_\ell$. Therefore uniqueness of y^* implies $y_r \in \Delta$ or $y_\ell \in \Delta$. The characterization using λ follows directly from the characterization of y^* in Lemma A7. □

Proposition A.3 (Effects of Extremism)**.** *Suppose Assumptions 1-2 and assume that ρ_E is fixed so that $\rho_R = 1 - \rho_L$. Increasing ρ_L:*

1. *weakly increases the locus of attraction, y^*; and*
2. *weakly increases the set of principals who strictly prefer a non-extremist, (E_L, E_R), in the strong set order sense.*

Proof. Fix $\rho_E \equiv \rho_L + \rho_R$. Thus, when we refer to increasing ρ_L throughout the proof, we are implicitly decreasing ρ_R by the same amount. Before proceeding, note that since ρ_E is constant, $A(y_d)$ is constant. Therefore $\frac{\partial \lambda(y_d)}{\partial \rho_L} - \frac{\partial \lambda(y_d)}{\partial \rho_R} = \frac{\partial \underline{x}(y_d)}{\partial y_d} + \frac{\partial \overline{x}(y_d)}{\partial y_d} > 0$ for all $y_d \in (y_\ell, y_r)$.

1. Since $\frac{\partial \lambda(y_d)}{\partial \rho_L} - \frac{\partial \lambda(y_d)}{\partial \rho_R} > 0$ for all $y_d \in (y_\ell, y_r)$, we know (i) $\lambda(y_\ell) \leq 0$ implies $y^* = y_\ell$, (ii) $\lambda(y_r) \geq 0$ implies $y^* = y_r$, and (iii) otherwise $\lambda(y^*) = 0$ at $y^* \in (y_\ell, y_r)$. Thus, y^* is weakly increasing.
2. From Lemma A8, E_L is the smallest $y_p < \underline{x}_\ell$ such that $\frac{\partial U(y_d, E_L)}{\partial y_d}\Big|_{y_d = \underline{x}_\ell^+} \geq 0$. Then, computation yields $\frac{\partial^2 U(y_d, E_L)}{\partial y_d \partial \rho_L}\Big|_{y_d = \underline{x}_\ell^+} - \frac{\partial^2 U(y_d, E_L)}{\partial y_d \partial \rho_R}\Big|_{y_d = \underline{x}_\ell^+} = -\frac{\delta \rho_d[(\underline{x}_\ell - E_L) + (\overline{x}_r - E_L)]}{1 - \delta \rho_E} < 0$, so E_L weakly increases in ρ_L. By an analogous argument, E_R weakly increases in ρ_L. □

D Competitive Representation

D.1 Proof of Lemma 5.

Direct computations show $\frac{\partial^2 \underline{x}(y_a, y_b)}{\partial y_a \partial y_b} < 0$. Using this fact, it is straightforward to sign $\frac{\partial y_a(y_b)}{\partial y_b}$ by applying the implicit function theorem to (6). The result for $y_b(y_a)$ is analogous. □

D.2 Proof of Remark 3.

Proposition 1 implies $y_a^* \in (y_{p_a}, y_m)$ and $y_b^* \in (y_m, y_{p_b})$. Lemma 5 implies uniqueness. □

D.3 Proof of Corollary 2.

The result follows directly from Lemma 5. If $\rho_L = \rho_R$, each principal's best response is independent of the opponent's choice and equals the single-principal optimum. Otherwise, the monotonicity of best responses implies the stated orderings. □

D.4 Proof of Proposition 2.

For an arbitrary profile of representative ideal points, $d = (d_1, d_2, ...d_n)$, let d_M denote the median when d is ordered by spatial location and let y_M denote the ideal point of the principal who appoints d_M. The first step in our proof establishes that in equilibrium, the median politician is appointed by the median principal, so $d_M^* = d_m^*$.

To prove this, we first show that $y_M \in \text{int}A^*$. We then use this to show that any politician to the left (right) of d_M^* must be appointed by a principal with an ideal point to the left (right) of y_M. Since d_M^* is the median politician, this implies that y_M is the median principal.

Proof that $y_M \in intA^*$: By Lemma 1, the equilibrium profile of politician ideal points d^* yields a unique equilibrium acceptance set, A^*, centered on d_M^*. By Proposition 1, given d_{-i}, it is optimal for principal i to appoint a median politician, $d_i \in [\ell_i, r_i]$, if and only if $y_i \in [\ell_i, r_i]$. Thus $d_M^* \in [\ell_M, r_M]$ and $y_M \in [\ell_M, r_M]$. Since all $d_M \in [\ell_M, r_M]$ would be the median politician, this implies that if principal M were to self appoint, $d_M = y_M$, then the acceptance set would be centered on his ideal point. Because both boundaries of A^* strictly increase in d_M on $[\ell_M, r_M]$ (Lemma 3), it follows that d_M^* must be located within $[\ell_M, r_M]$ such that $y_M \in \text{int}A^*$. ■

Proof that $d_i^* < (>)\, d_M^* \implies y_i < (>)\, y_M$: Suppose $d_i^* < d_M^*$, which implies $\ell_i = d_M^*$ and $\ell_M \geq d_i^*$. The latter further implies that $y_M \geq d_i^*$ since $y_M \in [\ell_M, r_M]$. First, if $d_i^* \in A^*$, then Proposition 1 yields that $d_i^* < \ell_i$ implies $y_i < d_i^*$. Since $y_M \geq \ell_M \geq d_i^*$, it follows that $y_i < y_M$. Otherwise, if $d_i^* \notin A^*$, then Proposition 1 yields that $d_i^* < \ell_i$ implies $y_i < \min A^*$. We know $y_M \in A^*$, so we must have $y_i < y_M$. An analogous proof establishes the result for $d_i^* > d_M^*$. ■

Since d_M^* is the median politician, it must be that y_M is the median principal. By substituting $M = m$ and applying contraposition to the second result from above, we obtain the second part of Proposition 2: $y_i > y_m \implies d_i^* \geq d_m^*$ and $y_i < y_m \implies d_i^* \leq d_m^*$. Finally, the third part then follows directly from Proposition 1. □

E Additional Extensions

E.1 Fixed Median

Consider the special case where $y_\ell = y_r$ (so y_m is constant) and $\rho_L = \rho_R$ (balanced extremist power). This setting provides additional tractability while isolating how representatives affect the acceptance set's radius rather than its center.

Proposition A.1 (Closed-Form Solution)**.** *Under Assumptions 1-2, if* $y_\ell = y_r$ *and* $\rho_L = \rho_R$*, then* $y_d^*(y_p)\big|_{(E_L,E_R)} = (1-\delta\rho_E)y_p + \delta\rho_E y_m$.

Proof. From Proposition A.1, if $y_\ell = y_r$ and $\rho_L = \rho_R$, then we have $y^* = y_m$ and symmetric acceptance sets. Direct computation using the first-order condition yields the result. ☐

Define the principal's value of representation as $\nu(y_p) \equiv U(y_d^*(y_p), y_p) - U(y_p, y_p)$.

Proposition A.2 (Value of Delegation)**.** *Suppose Assumptions 1-2. If* $y_\ell = y_r$ *and* $\rho_\ell = \rho_r = \rho_L = \rho_R$*, then:*

1. ν *is strictly increasing on* $[E_L, \underline{x}_\ell]$ *and strictly decreasing on* $[\underline{x}_\ell, y_m]$*; and*
2. ν *is strictly increasing on* $[y_m, \overline{x}_r]$ *and strictly decreasing on* $[\overline{x}_r, E_R]$.

Thus, the value of representation is maximized at the boundaries $\underline{x}_\ell$ *and* $\overline{x}_r$ *separating extremists from moderates.*

Proof. From Proposition A.1, $y_d^*(y_p)\big|_{(E_L,E_R)} = (1-\delta\rho_E)y_p + \delta\rho_E y_m$. Applying the envelope theorem yields $\nu'(y_p)\big|_{(E_L,\underline{x}_\ell)} = \rho_d(y_d^*(y_p) - \underline{x}_\ell) > 0$ and $\nu'(y_p)\big|_{(\underline{x}_\ell,y_m)} = \frac{-(\delta\rho_E)^2\rho_d(y_m-y_p)}{1-\delta\rho_E} < 0$. The result follows from continuity of ν. Analogously, ν strictly increases on $[y_m, \overline{x}_r]$ and strictly decreases on $[\overline{x}_r, E_R]$. ☐

E.2 Mass Representation

Consider $n \geq 3$ (odd) principals with ideal points $y_1 < y_2 < ... < y_n$ collectively choosing between two candidates with ideal points y_d and $y_{d'}$ via simple majority rule. All principals are policy-motivated with quadratic utility. After the representative is chosen, the n representatives bargain over policy as in the baseline model. We provide sufficient conditions for preferences to satisfy single-crossing, so that the median principal is decisive, and coalitions to be spatially ordered.

Proposition A.3 (Single-Crossing Property)**.** *If* $\max\{\tilde{P}(y_\ell), 1-\tilde{P}(y_\ell)\} \leq \frac{1}{2\delta}$*, then:*

1. $U(y_d, y_p)$ *satisfies the single-crossing condition on* $X \times X$*; and*
2. *The median principal is decisive under majority rule, with sufficiently right-leaning principals preferring the rightmost candidate in any pairwise comparison.*

Proof. It follows from Milgrom and Shannon (1994) and Cho and Duggan (2003) that it suffices to show $U(y_d, y_p)$ satisfies single-crossing on X. By Lemma 4, the stated condition implies $\mu(y_d)$ is increasing on X. Thus $U(y_d, y_p)$ satisfies increasing differences on X, which implies single-crossing. ☐

References

Alesina, Alberto and Howard Rosenthal, "A Theory of Divided Government," *Econometrica*, 1996, *64* (6), 1311–41.

Austen-Smith, David and Jeffrey Banks, "Elections, Coalitions, and Legislative Outcomes," *American Political Science Review*, 1988, *82* (2), 405–422.

_ and _ , "Monotonicity in Electoral Systems," *American Political Science Review*, 1991, *85* (2), 531–537.

Bailey, Michael A. and Matthew Spitzer, "Appointing Extremists," *American Law and Economics Review*, 2018, *20* (1), 105–137.

Banks, Jeffrey S. and John Duggan, "A Bargaining Model of Collective Choice," *American Political Science Review*, 2000, *94* (1), 73–88.

_ and _ , "A General Bargaining Model of Legislative Policy-Making," *Quarterly Journal of Political Science*, 2006, *1* (1), 49–85.

Baron, David P., "A Spatial Bargaining Theory of Government Formation in Parliamentary Systems," *American Political Science Review*, 1991, *85* (1), 137–164.

_ , "A Sequential Choice Theory Perspective on Legislative Organization," *Legislative Studies Quarterly*, 1994, *19* (2), 267–296.

_ , "A Dynamic Theory of Collective Goods Programs," *American Political Science Review*, 1996, *90* (2), 316–330.

_ and Ehud Kalai, "The Simplest Equilibrium of a Majority-Rule Division Game," *Journal of Economic Theory*, 1993, *61* (2), 290–301.

_ and John A. Ferejohn, "Bargaining in Legislatures," *American Political Science Review*, 1989, *83* (04), 1181–1206.

_ and John Ferejohn, "Bargaining and Agenda Formation in Legislatures," *American Economic Review*, 1987, *77* (2), 303–309.

Besley, Timothy and Stephen Coate, "Centralized Versus Decentralized Provision of Local Public Goods: A Political Economy Approach," *Journal of Public Economics*, 2003, *87* (12), 2611–2637.

Buisseret, Peter and Dan Bernhardt, "Dynamics of Policymaking: Stepping Back to Leap Forward, Stepping Forward to Keep Back," *American Journal of Political Science*, 2017, *61* (4), 820–835.

Bulow, Jeremy I., John D. Geanakoplos, and Paul D. Klemperer, "Multimarket Oligopoly: Strategic Substitutes and Complements," *Journal of Political Economy*, 1985, *93* (3), 488–511.

Cardona, Daniel and Clara Ponsati, "Uniqueness of Stationary Equilibria in Bargaining One-Dimensional Policies Under (Super) Majority Rules," *Games and Economic Behavior*, 2011, *73* (1), 65–75.

Cho, Seok-Ju and John Duggan, "Uniqueness of Stationary Equilibria in a One-Dimensional Model of Bargaining," *Journal of Economic Theory*, 2003, *113* (1), 118–130.

Christiansen, Nels, "Strategic Delegation in a Legislative Bargaining Model with Pork and Public Goods," *Journal of Public Economics*, 2013, *97* (C), 217–229.

Cox, Gary W. and Mathew D. McCubbins, *Legislative Leviathan: Party Government in the House*, Cambridge University Press, 2007.

Diermeier, Daniel and Razvan Vlaicu, "Parties, Coalitions, and the Internal Organization of Legislatures," *American Political Science Review*, 2011, *105* (2), 359–380.

_ and Roger B. Myerson, "Bicameralism and Its Consequences for the Internal Organization of Legislatures," *American Economic Review*, 1999, *89* (5), 1182–1196.

_ , Carlo Prato, and Razvan Vlaicu, "Procedural Choice in Majoritarian Organizations," *American Journal of Political Science*, 2015, *59* (4), 866–879.

_ , _ , and _ , "A Bargaining Model of Endogenous Procedures," *Social Choice and Welfare*, 2016, *47* (4), 985–1012.

_ , _ , and _ , "Self-enforcing Partisan Procedures," *Journal of Politics*, 2020, *82* (3), 937–954.

Dixit, Avinash, "The Role of Investment in Entry-Deterrence," *The Economic Journal*, 1980, *90* (357), 95–106.

Duch, Raymond M., Jeff May, and Dave Armstrong, "Coalition-Directed Voting in Multi-Party Democracies," *American Political Science Review*, 2010, *104* (4), 698–719.

Duggan, John, "Majority Voting over Lotteries: Conditions for Existence of a Decisive Voter," *Economics Bulletin*, 2014, *34* (1), 263–270.

_ and Mark Fey, "Repeated Downsian Electoral Competition," *International Journal of Game Theory*, 2006, *35* (1), 39–69.

_ and Tasos Kalandrakis, "Dynamic Legislative Policy Making," *Journal of Economic Theory*, 2012, *147* (5), 1653–1688.

Eguia, Jon X. and Kenneth A. Shepsle, "Legislative Bargaining with Endogenous Rules," *Journal of Politics*, 2015, *77* (4), 1076–1088.

Epstein, David and Sharyn O'Halloran, "Legislative Organization under Separate Powers," *Journal of Law, Economics, and Organization*, 2001, *17* (2), 373–396.

Eraslan, Hülya and Kirill S Evdokimov, "Legislative and multilateral bargaining," *Annual Review of Economics*, 2019, *11* (1), 443–472.

Eulau, Heinz and Paul D. Karps, "The Puzzle of Representation: Specifying Components of Responsiveness," *Legislative Studies Quarterly*, 1977, *2* (3), 233–254.

Evans, C. Lawrence, "Congressional Committees," in George C. Edwards III, Frances E. Lee, and Eric Schickler, eds., *The Oxford Handbook of the American Congress*, Oxford University Press, 2011, pp. 29–56.

Fenno, Richard, *Congressmen in Committees*, Boston, MA: Little, Brown, 1974.

Fershtman, Chaim, Kenneth L. Judd, and Ehud Kalai, "Observable Contracts: Strategic Delegation and Cooperation," *International Economic Review*, 1991, *32* (3), 551–559.

Fowler, James H., "Elections and Markets: The Effect of Partisanship, Policy Risk, and Electoral Margins on the Economy," *Journal of Politics*, 2006, *68* (1), 89–103.

Friedrich, Carl J., *Constitutional Government and Politics*, Harper: New York, 1937.

Gailmard, Sean and Thomas Hammond, "Intercameral Bargaining and Intracameral Organization in Legislatures," *Journal of Politics*, 2011, *73* (2), 535–546.

Hall, Richard L. and Bernard Grofman, "The Committee Assignment Process and the Conditional Nature of Committee Bias," *American Political Science Review*, 1990, *84* (4), 1149–1166.

Harstad, Bård, "Strategic Delegation and Voting Rules," *Journal of Public Economics*, 2010, *94* (1), 102–113.

Judd, Gleason, "Access to Proposers and Influence in Collective Policy Making," *Journal of Politics*, 2023, *85* (4), 1430–1443.

Kalandrakis, Tasos, "Proposal Rights and Political Power," *American Journal of Political Science*, 2006, *50* (2), 441–448.

Kang, Myunghoon, "Representation, Sophisticated Voting, and the Size of the Gridlock Region," *Journal of Theoretical Politics*, 2017, *29* (4), 623–646.

Kedar, Orit, "When Moderate Voters Prefer Extreme Parties: Policy Balancing in Parliamentary Elections," *American Political Science Review*, 2005, *99* (2), 185–199.

_ , *Voting for Policy, Not Parties: How Voters Compensate for Power Sharing*, Cambridge University Press, 2009.

Kenderov, P., "Multivalued Monotone Mappings are Almost Everywhere Single-valued," *Studia Mathematica*, 1976, *56* (3), 199–203.

Klumpp, Tilman, "Strategic Voting and Conservatism in Legislative Elections," *working paper*, 2010, (Available at: `https://sites.ualberta.ca/~klumpp/docs/representation.pdf`).

Krasa, Stefan and Mattias K. Polborn, "Political Competition in Legislative Elections," *American Political Science Review*, 2018, *112* (4), 809–825.

Krehbiel, Keith, "Are Congressional Committees Composed of Preference Outliers?," *American Political Science Review*, 1990, *84* (1), 149–163.

_ , *Pivotal Politics: A Theory of US Lawmaking*, University of Chicago Press, 1998.

Loeper, Antoine, "Cross-border Externalities and Cooperation Among Representative Democracies," *European Economic Review*, 2017, *91* (C), 180–208.

McCarty, Nolan, "Proposal Rights, Veto Rights, and Political Bargaining," *American Journal of Political Science*, 2000, *44* (3), 506–522.

Miettinen, Topi, "Commitment Tactics in Bargaining under Complete Information," in Emin Karagozoglu and Kyle B. Hyndman, eds., *Bargaining: Current Research and Future Directions*, Springer, 2022, pp. 11–34.

Milgrom, Paul and Chris Shannon, "Monotone Comparative Statics," *Econometrica*, 1994, *62* (1), 157–180.

Miller, Warren E. and Donald E. Stokes, "Constituency Influence in Congress," *American Political Science Review*, 1963, *57* (1), 45–56.

Patty, John W. and Elizabeth Maggie Penn, "Are Moderates Better Representatives than Extremists? A Theory of Indirect Representation," *American Political Science Review*, 2019, *113* (3), 743–761.

Persson, Torsten and Guido Tabellini, "The Politics of 1992: Fiscal Policy and European Integration," *Review of Economic Studies*, 1992, *59* (4), 689–701.

Pitkin, Hanna F., *The Concept of Representation*, University of California Press, 1967.

Rohde, David W. and Kenneth A. Shepsle, "Democratic Committee Assignments in the House of Representatives: Strategic Aspects of a Social Choice Process," *American Political Science Review*, 1973, *67* (3), 889–905.

Romer, Thomas and Howard Rosenthal, "Political Resource Allocation, Controlled Agendas, and the Status Quo," *Public Choice*, 1978, *33* (4), 27–43.

Schelling, Thomas C., "An Essay on Bargaining," *American Economic Review*, 1956, *46* (3), 281–306.

Segendorff, Björn, "Delegation and Threat in Bargaining," *Games and Economic Behavior*, 1998, *23* (2), 266–283.

Simon, Herbert A., "Notes on the Observation and Measurement of Political Power," *Journal of Politics*, 1953, *15* (4), 500–516.

Sobel, Joel, "Distortion of Utilities and the Bargaining Problem," *Econometrica*, 1981, *49* (3), 597–619.

Vickers, John, "Delegation and the Theory of the Firm," *The Economic Journal*, 1985, *95*, 138–147.

Volden, Craig, "A Formal Model of the Politics of Delegation in a Separation of Powers System," *American Journal of Political Science*, 2002, *46* (1), 111–133.

Zápal, Jan, "Simple Markovian Equilibria in Dynamic Spatial Legislative Bargaining," *European Journal of Political Economy*, 2020, *63* (C), 101816.