Representation in Collective Policymaking*

Daniel Gibbs[†] Gleason Judd[‡]

Abstract

Representatives often participate in collective policymaking. What are their effects on policy and who do principals want as their representative? We model a representative who bargains with other politicians to set policy. The representative's ideology influences their own behavior, which shifts everyone's expectations about policymaking and changes what can pass, along with extremist proposals. To induce extremist moderation, many principals want to improve the policymaking expectations of centrist (de facto) veto players by choosing a representative biased towards them. This preference for more centrist representatives and against more extreme representatives is general and widespread. Under standard assumptions, principals bias towards a unique central location, which varies with the balance of extremist proposal rights. In extensions, we study collective representation, the value of representation, and competitive representation. Our results shed light on representation in separation-of-powers systems and congressional committees, as well as elections into those bodies.

*We thank audiences at Caltech, Princeton, Washington University in St. Louis, Virginia Tech, the 2022 BEET workshop, and the 2023 CEPR/Warwick/Princeton/Yale Political Economy Symposium for comments and suggestions.

[†]Virginia Tech

[‡]Princeton

Two core features of modern democracies are representatives and collective policymaking. Frequently, they are combined. For example, legislators represent party leaders in committees, and election-winners represent voters in legislatures or separation-of-powers systems. In these contexts, representatives can have subtle effects on policymaking that obscure their influence and appeal (Miller and Stokes 1963; Pitkin 1967; Eulau and Karps 1977),[1] so evaluating and comparing them requires well-developed theory (Gailmard and Hammond 2011). To build our theoretical understanding, we revisit two fundamental questions about representation in collective bodies. First, how do different representatives affect collective policymaking and, in turn, the welfare of those being represented? Second, who does the represented want as their representative?

We study a game-theoretic model in which a representative bargains with other politicians by proposing and voting over one-dimensional policy until majority agreement. Players propose strategically whenever given the opportunity and following any proposal a majoritarian vote determines whether it passes or bargaining continues. In equilibrium, whoever is recognized first will propose and pass their favorite majority-approved policy. Those passable policies coincide with the median politician's *acceptance set*, an interval of policies around her ideal point that is uniquely determined by her vote calculus. Crucially, since policymaking can continue after rejected proposals, that vote calculus depends on expectations about continued policymaking — which depend on the full profile of politician ideal points and their proposal rights. Thus, the representative's ideal point can affect what passes through two channels: the median's location and expectations. In turn, it can alter proposals by other politicians.

We emphasize an underappreciated mechanism for the representative's influence on policymaking: her ideology can *shift policymaking expectations*. Although this force is present in the legislative bargaining models we build on (e.g, Banks and Duggan 2000, 2006) and also

[1] Miller and Stokes (1963) highlighted 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) echoed that "[. . .] representatives are influenced in their conduct by many forces or pressures or linkages [. . .]" (pg. 235). More broadly, Pitkin (1967) stated 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).

related to the classic idea of *anticipation effects* (Friedrich 1937; Simon 1953), its consequences for representation are undeveloped. We show how it can present principals with the familiar tradeoff that a biased representative may induce other politicians to propose more favorable policies (Schelling 1956). Moreover, we analyze how it interacts with the more familiar mechanism of shifting the median's location (Klumpp 2010; Gailmard and Hammond 2011).

We show that this pervasive mechanism encourages bias towards (de facto) veto players. Merely shifting the representative in their direction will improve their expectations, narrow the acceptance set, and thereby constrain extreme proposals. This force encourages moderation — the principal always wants someone who is centrist or biased in that direction, never someone more extreme. Yet, it has limits — sufficiently extreme principals do not want a centrist.

Our most general results characterize optimal representatives for different principals. Centrist principals always want nearby representatives who will be the median, but may bias either way. Moderate principals want their representative to be strictly more centrist, potentially enough to be the median. The downside of her representative's farther proposal is outweighed by the upside of improving the median's policymaking expectations and further constraining extremists on both sides. Extremist principals want a representative who is on their side of the centrists, potentially moderate or extreme. Finally, neither extremists nor extreme moderates want their representative to be the median.

Additionally, in a setting where players have quadratic policy utility and are (mostly) polarized, we fully characterize preferences over representatives. Unless the principal is very extreme, she has a unique optimal representative who is biased inwards towards a unique centrist location. This *locus of attraction* characterizes the unique principal who strictly prefers to have an unbiased representative.[2] Moreover, optimal representatives are ordered. Thus, (i) the principals who want a centrist representative are an interval containing the centrists along with some moderates on at least one side, and (ii) around that interval are the two intervals of principals who want a moderate representative. Furthermore, there is

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

always a *dead zone* of centrist and moderate representatives who are *not* optimal for *any* principal. Finally, we characterize how the balance of extremist proposal rights affects these preferences and show that principals skew away from the gaining side in order to further constrain those extremists.

To enrich the application and interpretation of our results, we explore three extensions. First, we study *collective representation* — i.e., collective choice over the representative's ideal point. For instance, majoritarian collective choice coincides with the median principal's preference under broad conditions, since preferences over representatives satisfy single-crossing as long as extremist proposal rights are not too high. Second, we study the *value of representation* — i.e., the principal's welfare gain from optimal representation. This value is greatest for principals near each moderate-extremist boundary if extremist proposal rights are balanced. Third, we study *competitive representation* — i.e., two opposing principals each choosing a representative in a setting with two open positions. Both principals moderate their representative towards the median, which can encourage or discourage further moderation depending on the balance of extremist proposal rights.

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).[3] For example, a status-quo-biased representative who is a veto player will be more demanding and thereby further constrain extreme proposals (Gailmard and Hammond 2011; Klumpp 2010). We find a similar incentive to constrain extremists by making veto players more demanding, but through a different mechanism — the representative's effect on expectations about policymaking — that is lurking in well-known models of collective policymaking (Banks and Duggan 2000, 2006). Since both mechanisms may be present in various settings (e.g., Banks and Duggan 2006), our results complement earlier work by showing a force for

[3] 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).

moderation that 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 be appealing by reducing extremism. Thus, we add to theoretical understanding of how incentives to counteract extremism can encourage moderation in various aspects of collective policymaking. 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).[4] In contrast, we fix (possibly unequal) proposal rights and show a widespread preference for relatively centrist representatives — to make extreme proposals *less extreme*.[5] Second, during bargaining that can continue with accepted policy as the new status quo, proposers may opt for relatively centrist policy that directly increases the median's reservation value in *future* periods and thus constrains their 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*. Furthermore, in our analysis moderate principals want to constrain extremists on both sides, not just their opponents. Third, when acquiring access to facilitate lobbying on policy proposals, interest groups prefer to target more extreme representatives in order to increase the chances of moderating their proposals, thereby also improving centrist expectations and constraining what extremists can pass (Judd 2023). Our results highlight that beforehand, when the representatives are chosen, similar incentives encourage those groups to support the selection of more moderate candidates.

The moderation incentives we uncover also complement extremism incentives driven by collective policymaking in other settings. In a supermajoritarian take-it-or-leave-it setting, voters never want someone more moderate but may prefer a strictly more extreme represen-

[4]For other related work on *endogenous procedures*, see Diermeier and Vlaicu (2011); Diermeier et al. (2015, 2016). In 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.

[5]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.

tative who would be a veto pivot (Kang 2017). In other contexts where policy is a weighted average of politician ideal points, extreme representatives can counterbalance extreme opponents (Alesina and Rosenthal 1996; Kedar 2005, 2009). Additionally, if principals care about how their representative will vote on an *exogenous* legislative agenda, then preferences can be asymmetric and favor extremism (Patty and Penn 2019). Understanding these different directions can help inform empirical interpretation and anticipation of future choices.[6]

We provide implications for understanding representation and representatives in various contexts. Our model of collective policymaking is a *minimal legislative process* (Baron 1994) that allows several interpretations[7] and provides a lens for studying representation in separation-of-powers systems (Epstein and O'Halloran 2001; Volden 2002) or, more narrowly, the *representativeness* of congressional committees (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).[8] Broadly, we emphasize *partisan* factors and complement related work emphasizing *distributive* and *informational* factors.[9] More precisely, we probe the role of the broader partisan environment[10] and find that (i) some centrists and extremists will want unbiased representatives, (ii) party leaders like moderation unless they are very extreme, and (iii) polarization strengthens the desire for moderation, especially on the weaker side. Furthermore, the *dead zone* provides a logic for why key committees may have centrist and extremist members, but none in between. Those intermediate legislators are not optimal representatives for anyone; neither themselves nor party leaders.

[6]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).

[7]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).

[8]For an overview of scholarship in committee composition, see Evans (2011).

[9]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).

[10]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).

Additionally, our collective representation extension provides a foundation for new insights into mass representation, including: (i) voting in elections for collective policymaking positions (Kedar 2005, 2009; Duch et al. 2010), (ii) electoral competition over those offices (Austen-Smith and Banks 1988; Krasa and Polborn 2018), and (iii) its representativeness (Austen-Smith and Banks 1991). We shed new light on how expectations about collective policymaking can affect incentives of party leaders and voters (Kedar 2005, 2009; Duch et al. 2010), thereby influencing who gets nominated and their electoral chances. For instance, we show that unless extremist proposal rights are very high, electoral competition in which each party chooses a candidate representative will feature a unique indifferent voter. Thus, (i) each candidate's win probability is easy to characterize, and (ii) both parties will converge towards the median voter's optimal representative.

Model

Players. There is a principal, P; a continuum of potential representatives; and a finite set K of $k \geq 2$ (even) legislators in office.

Timing. The model has two stages. First, in the *appointment* stage, P selects a representative, denoted d, to bargain on her behalf.[11] Second, in the *bargaining* stage, the representative d interacts with the legislators in K to collectively select a policy within a one-dimensional policy space, $X = [0, 1]$. 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 $x^t \in X$. 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 by a simple majority of politicians, 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.

[11] We follow convention in the principal-agent literature by referring to a legislator (agent) as *he* and the principal as *her*.

Preferences. All players are purely policy-motivated and each player has a unique ideal point $y_i \in X$. More precisely, i's policy utility from a policy x is $u(x, y_i)$, where u is continuously differentiable, strictly concave in x, symmetric around a single peak at y_i, and identical up to translations in y_i.[12] The ideal points of the k legislators in K are ordered such that $y_1 \leq y_2 \leq ... \leq y_k$, and we denote $l = \frac{k}{2}$ and $r = \frac{k}{2} + 1$ for convenience. We denote the principal's ideal point as y_p and the ideal point of her chosen representative as y_d.

We study a *bad status quo* setting in which (i) all agents receive zero utility during each period until agreement and (ii) enacting policy x gives utility $u(x, y) \geq 0$ to a player with ideal point y. This setting clearly highlights the role of policymaking expectations and sharpens our results. Our key forces and insights are robust to including a status quo policy.

Cumulative payoffs are sums of per-period utilities, discounted by the common factor $\delta \in (0, 1)$. For convenience, we normalize per-period utility by the factor $1 - \delta$. Thus, if x is accepted in period t, then legislator 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 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.[13] A *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 a stationary subgame perfect equilibrium and (ii) P chooses y_d to maximize her expected payoff anticipating the distribution of policy outcomes that y_d will induce.

[12] Specifically, a strictly decreasing, concave, and continuously differentiable function $g : [0, 1] \to \mathbb{R}_+$ exists such that $u(x, y) = g(|x - y|)$.

[13] 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).

Analysis

To begin, we characterize equilibrium behavior during the bargaining stage, after y_d is chosen. Then, we trace how y_d affects d's behavior, as well as 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 extensions.

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).

Specifically, each politician always prefers to propose the closest majority-approved policy rather than delay, so equilibrium policymaking is characterized by the *acceptance set* of policies that would pass if proposed.[14] The acceptance set is unique and determined entirely by the voting calculus of the politician with the median ideal point. That politician, whom we denote m, will pass any proposal that she prefers over her continuation value from rejecting to continue bargaining, denoted V_m. Thus, the acceptance set is a compact interval in X that is centered at m's ideal point, denoted y_m, with a radius that depends on V_m.

The acceptance set can vary with y_d through y_m or V_m. First, m's ideal point is:

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

Second, y_d can affect m's continuation value by shifting y_m or by shifting d's proposal.

[14] These properties follow immediately from Cardona and Ponsati (2011).

Specifically, given an acceptance interval $[\underline{x}, \overline{x}]$, that value is:

$$V_m(\underline{x}, \overline{x}) \equiv P(\underline{x})u(\underline{x}, y_m) + (1 - P(\overline{x}))u(\overline{x}, y_m) + \sum_{i \in N : y_i \in (\underline{x}, \overline{x}]} \rho_i u(y_i, y_m), \tag{2}$$

where $P(x) \equiv \sum_{i \in N : y_i \leq x} \rho_i$ denotes the cumulative proposal rights of politicians left of x. In (2), the first term is m's utility from the lower bound $\underline{x}$ weighted by the proposal rights of politicians with $y_i \leq \underline{x}$, the second term is analogous for the upper bound $\overline{x}$, and the last term is the proposal-rights-weighted sum of m's utility from $y_i \in (\underline{x}, \overline{x}]$.

The equilibrium acceptance set is $A(y_d) = [\underline{x}_m(y_d), \overline{x}_m(y_d)]$, where

$$\underline{x}_m(y_d) = \min \{x \in X \mid x < y_m \text{ and } u(x, y_m) \geq \delta V_m(\underline{x}_m(y_d), \overline{x}_m(y_d))\}$$

is the leftmost proposal that would pass, and $\overline{x}_m(y_d)$ is analogous for $x > y_m$. Thus, it is consistent with each politician i's equilibrium proposal strategy specifying the unique acceptable policy closest to y_i. Furthermore, the connection between proposals and voting flows through V_m.

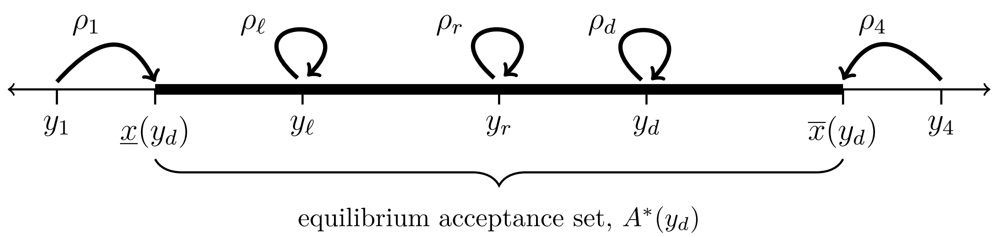
Lemma 1 summarizes properties of equilibrium policymaking given y_d.

Lemma 1 (Cardona and Ponsati (2011))**.** *For each $y_d \in X$, the bargaining subgame has a unique stationary subgame perfect equilibrium, which has acceptance set $A(y_d) = [\underline{x}_m(y_d), \overline{x}_m(y_d)]$; and each politician i proposes the $x \in A(y_d)$ minimizing $|x - y_i|$.*

Lemma 1 implies that every y_d induces a unique equilibrium policy lottery. Essentially, the boundaries of $A(y_d)$ and the ideal points in its interior are each weighted by the recognition probability of the politicians who propose them. A key implication is that every potential representative induces well-defined expected payoffs.

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

Figure 1: Illustration of equilibrium policymaking (given y_d)



Note: Figure 1 illustrates Lemma 1 for a hypothetical legislature with $K = 4$ and $y_d > y_r$, which implies $y_m = y_r$. The acceptance set is the bold interval. Arrows point from legislators to their proposals (if recognized). Each legislator proposes the closest acceptable policy.

To streamline presentation, we maintain two mild assumptions ensuring at least one politician is always constrained. Essentially, we suppose players are not too impatient and some politician is sufficiently extreme. First, δ is high enough that the most extreme policies will not pass regardless of y_d (Remark 2, Assumption 1). Intuitively, m is quite willing to wait for better proposals and — although y_d can affect that willingness — never settles for accepting the most extreme policies. Second, at least one of the most extreme politicians is always outside of the equilibrium acceptance set regardless of y_d (Assumption 2).

Remark 2. *There exists $\overline{\delta} \in (0, 1)$ such that $\delta > \overline{\delta}$ implies $A(y_d) \subset intX$ for all $y_d \in X$.*[15]

Assumption 1. *Suppose $\delta > \overline{\delta}$.*

Assumption 2. *For all $y_d \in X$, we have $\{y_1, y_k\} \not\subset A^*(y_d)$.*

The representative's effects on policymaking

We now study how the representative's ideal point, y_d, affects policymaking. First, we characterize when it affects the representative's proposals. Then, we show how it affects the acceptance set and, in turn, proposals by other politicians.

The representative's proposal varies with y_d if and only if she will not be constrained by the acceptance set — i.e., $y_d \in \text{int } A(y_d)$. Lemma 2 shows that this property is fully characterized by a centrally-located open interval.

[15] This observation follows from (Banks and Duggan 2000).

Lemma 2. *There are unique* $\underline{x}_\ell \in (0, y_\ell)$ *and* $\overline{x}_r \in (y_r, 1)$ *such that* $y_d \in intA^*(y_d)$ *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 representatives based on whether they would be the median politician or propose a boundary of the acceptance set. First, $y_d \in [y_\ell, y_r]$ would be the median and propose their (interior) ideal point. Second, $y_d \notin (\underline{x}_\ell, \overline{x}_r)$ would not be the median and would propose the nearest boundary of the acceptance set. Finally, $y_d \in (\underline{x}_\ell, y_\ell) \cup (y_r, \overline{x}_r)$ would not be the median but would propose their (interior) 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.*

In equilibrium, y_d can affect the acceptance set by shifting its center, y_m, or radius, via V_m, depending on the location of y_d. Over centrists, y_d affects both — since $y_m = y_d$ and therefore y_m shifts relative to the other potential proposers, so V_m can also vary. Over aligned moderates, y_d affects only the radius — since y_m is constant but V_m varies due to changes in d's proposal. Thus, the representative can affect the acceptance set even without shifting the median. Finally, over aligned extremists, y_d does not affect either — since y_m is constant, as is V_m because these representatives all propose the same boundary.

Broadly, the acceptance set is (i) constant in y_d over aligned extremists, (ii) shrinks inward as y_d shifts inward over aligned moderates, and (iii) shifts in the same direction as y_d over centrists. For (i), y_m and V_m are constant over aligned extremists. For (ii), V_m increases as y_d approaches y_m over aligned moderates, so both boundaries shift inward and all aligned moderates generate distinct acceptance sets. Yet, aligned moderates near their median will induce nearly identical acceptance sets — since y_m is constant and the effect through V_m vanishes as y_d approaches y_m because u is strictly concave and differentiable. Finally, for (iii), the effects through y_m and V_m can oppose each other but the first always dominates, so both boundaries shift the same direction as y_d and all centrists generate distinct acceptance sets. Lemma 3 makes these observations precise.

Lemma 3. *The correspondence A^* is continuous and $A^*(y_d) \subset [\underline{x}_\ell, \overline{x}_r]$ for all y_d, with:*

1. *$y_d \leq \underline{x}_\ell$ implies $A^*(y_d) = [\underline{x}_\ell, \overline{x}_\ell]$;*

2. *$y_d \in (\underline{x}_\ell, y_\ell)$ implies $A^*(y_d) = [\underline{x}_\ell(y_d), \overline{x}_\ell(y_d)] \subset [\underline{x}_\ell, \overline{x}_\ell]$, where $\underline{x}_\ell(y_d)$ strictly increases and $\overline{x}_\ell(y_d)$ strictly decreases, each at rates that converge to zero as $y_d \to y_\ell$;*

3. *$y_d \in [y_l, y_r]$ implies $A^*(y_d) = [\underline{x}_d(y_d), \overline{x}_d(y_d)] \subset [\underline{x}_\ell(y_\ell), \overline{x}_r(y_r)]$, where $\underline{x}_d(y_d)$ and $\overline{x}_d(y_d)$ strictly increase;*

4. *$y_d \in (y_r, \overline{x}_r)$ implies $A^*(y_d) = [\underline{x}_r(y_d), \overline{x}_r(y_d)] \subset [\underline{x}_r, \overline{x}_r]$, where $\underline{x}_r(y_d)$ strictly decreases and $\overline{x}_r(y_d)$ strictly increases, each at rates that converge to zero as $y_d \to y_r$; and*

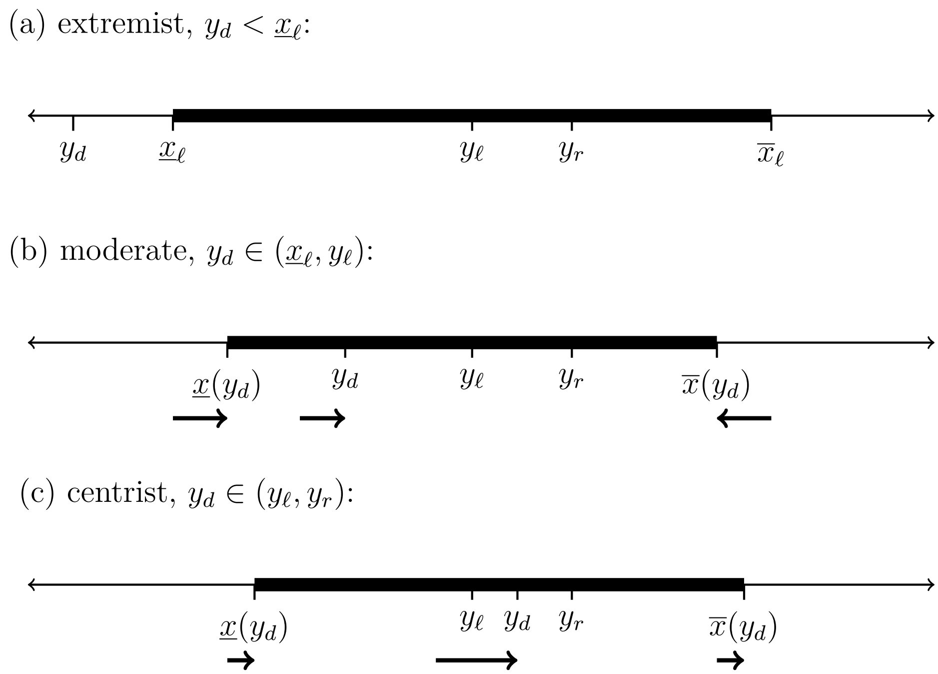
5. *$y_d \geq \overline{x}_r$ implies $A^*(y_d) = [\underline{x}_r, \overline{x}_r]$.*

Lemma 3 has several additional implications for how the acceptance set varies with y_d. First, extremist representatives induce a larger acceptance set than their aligned moderates and, moreover, it is more extreme on both ends — e.g., $y_d \in (\underline{x}_\ell, y_\ell]$ implies $[\underline{x}(y_d), \overline{x}(y_d)] \subset [\underline{x}_\ell, \overline{x}_\ell]$. Second, extremist principals are always outside the acceptance set but are closest if d is an aligned extremist — i.e., the minimum lower bound is $\underline{x}_\ell$ and the maximum upper bound is $\overline{x}_r$.[16]

The policymaking effects of y_d depend on its location, as illustrated in Figure 2. First, aligned extremists are equivalent since they induce the same acceptance set and propose the same boundary (see Figure 2(a)). In contrast, opposing extremists are distinct since they propose differently and may also alter other proposals. Next, non-extremists are also distinct. Among moderates, those closer to the center shift all proposals weakly inwards by (i) proposing more centrist policy and (ii) inducing smaller acceptance sets (see Figure 2(b)). And among centrists, those who lean farther right shift every proposal weakly rightward by (i) proposing more right-leaning policy and (ii) shifting the acceptance set to the right (see Figure 2(c)).

[16] See Lemma A1.

Figure 2: How the representative affects policymaking



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 shifts 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 $y_d > y_r$.

Optimal Representation

We apply our understanding of the representative's various policymaking effects in order to provide our most general results about optimal representatives. We show that the principal wants a representative who is centrist or biased that direction, and never wants someone more extreme. Moreover, we shed light on the limits of moderation: sufficiently extreme principals do not want a centrist representative.

To do so, we study how y_d affects P's equilibrium value from bargaining. That value is always uniquely defined since the equilibrium policy lottery is unique, so for each pair of y_d and y_p we denote it $U(y_d, y_p)$. Furthermore, the equilibrium characterization implies

$$U(y_d, y_p) = V_p(\underline{x}(y_d), \overline{x}(y_d)). \tag{3}$$

We analyze the set of y_d that maximize P's equilibrium value and how it varies with y_p. Formally, 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.

Although equilibrium strategies are unique during bargaining, P can have multiple optimal representatives — i.e., y_d^* is not necessarily single valued. Different representatives will shift and shape A^* in different ways. Even if two different representatives induce different policymaking, they may be equivalent for P by trading off worse representative proposals against better extremist proposals, or vice versa. The two key considerations are (i) whether P wants to constrain both extremists, and (ii) if not, which extremist they want to constrain.

First, we narrow the scope for locating optimal representatives by shedding light on who is *not* optimal. Specifically, Lemma 4 shows that P never wants a representative from the other side of the spectrum.

Lemma 4. *If $y_p \leq y_r$, then $y^*(y_d) \subset [0, y_r]$. If $y_p \geq y_\ell$, then $y^*(y_d) \subset [y_\ell, 1]$.*

Lemma 4 is natural but the logic is not trivial. Although opposing representatives have clear downside for P since their proposal will be farther away and they relax the constraint on opposing extremists, they can also have some upside since they relax the constraint on the aligned extremists. Yet, they are always strictly dominated by some aligned representative who equally constrains the aligned extremists but provides less downside from other proposals.[17]

Next, we characterize which representatives are optimal for different principals. Specifically, Proposition 1 establishes properties of y_d^* that bound the set of optimal representatives for each y_p. We show: (i) centrist principals want only centrist representatives, (ii) moderates want either a centrist or more-centrist aligned moderate, and (iii) extremists want extremists or aligned moderates. Moreover, extremists and sufficiently extreme moderates do not want a centrist.

[17]See Lemma A2.

Proposition 1 (Optimal representatives)**.** *The optimal representative correspondence y_d^* is non-empty, upper hemicontinuous, and compact valued. Under Assumptions 1 and 2:*

1. *$y_p \in [y_\ell, y_r]$ implies $y_d^*(y_p) \subseteq [y_\ell, y_r]$;*

2. *$y_p \in (\underline{x}_\ell(y_\ell), y_\ell)$ implies $y_d^*(y_p) \subset (y_p, y_r] \backslash \{y_\ell\}$, and analogously for $y_p \in (y_r, \overline{x}_r(y_r))$;*

3. *$y_p \in (\underline{x}_\ell, \underline{x}_\ell(y_\ell)]$ implies $y^*(y_p) \subset (y_p, y_\ell)$, and analogously for $y_p \in [\overline{x}_r(y_r), \overline{x}_r)$; and*

4. *$y_p \leq \underline{x}_\ell$ implies $y_d^*(y_p) \subset [0, y_\ell)$ and $y_d^*(y_p) \cap [\underline{x}_\ell, y_\ell) \neq \emptyset$, and analogously for $y_p \geq \overline{x}_r$.*

Proposition 1 shows that P may want to moderate but never strictly prefers a more extreme representative. More precisely, optimal representatives are centrist or biased in that direction for a broad set of principals that includes centrists, moderates, and potentially some extremists. First, centrist principals always want centrist representatives — some centrist can always improve upon the only appeal of any non-centrist representative, as in the discussion following Lemma 4. Second, moderate principals prefer some moderation — a negligible loss in their representative's proposal comes with a tangible gain from moderating extreme proposals. Finally, extremism is never appealing — a more extreme representative is never better than an ally, and for centrist and moderate principals they are strictly worse. Thus, only an extremist principal may want an extreme representative.

Additionally, Proposition 1 reveals limits to moderation incentives. First, some extremist principals may not want to moderate at all. Second, sufficiently extreme principals do not want a centrist representative. For extremists, (i) their nearest centrist is always worse than nearly-centrist aligned moderates, since the acceptance set is approximately equal; and (ii) all farther centrists are even worse, since the acceptance set shifts away. Similar forces make centrist representatives unappealing to moderate principals who are sufficiently extreme.

Next, Corollary 1.1 highlights that centrist and extremist principals may want an ally representative but moderates never do.

Corollary 1.1. *In equilibrium, (i) a centrist principal may strictly prefer an ally representative, (ii) a moderate never wants an ally, and (iii) an extremist may weakly prefer an ally but never strictly prefers one.*

An ally representative is always best if they propose, so a non-ally is optimal only if they outweigh the downside of their own proposal by making extremist proposals sufficiently more favorable for P. First, moderate principals never want an ally representative. They always gain by moderating slightly to constrain extremists on both sides, as discussed above. Second, extremists can weakly prefer their ally (along with all other aligned extremists). They face a non-negligible proposal cost from moderating enough to affect the acceptance set and their return from constraining extremists always depends on relative extremist proposal power. Finally, centrists can strictly prefer their ally. The negligible proposal cost of bias in either direction can coincide with a negligible effect on extremist proposals, since one shifts closer and the other shifts away.

We can also shed light on how optimal representatives are ordered and which direction P wants to bias their representative. For centrist principals, there is a clear ordering — y_d^* is an increasing correspondence over $[y_\ell, y_r]$[18] — but P may want to bias in either direction. In contrast, for both moderates and extremists the ordering is unclear — since the policy lottery induced by y_p is not naturally ordered on X[19] — but their preferred bias is clear. Moderates want to bias strictly towards their opposite pivot. Extremists want to bias towards their closest centrist whenever an ally is not optimal.

Finally, Corollary 1.2 sharpens the characterization for the fixed median case, i.e., $y_\ell = y_r$.

Corollary 1.2. *If $y_\ell = y_r = y_m$, then (i) $y_p = y_m$ implies $y_d^*(y_p) = y_m$, (ii) $y_p \in (\underline{x}_\ell, y_m)$ implies $y_d^*(y_p) \subset (y_p, y_m)$, and (iii) $y_p < \underline{x}_\ell$ implies $y_d^*(y_p) < y_m$.*

[18] Specifically, U has the single-crossing property on $[y_\ell, y_r] \times X$ because equilibrium proposals are increasing over $y_d \in [y_\ell, y_r]$ and u satisfies the single-crossing property. Then, since $y_p \in [y_\ell, y_r]$ implies $y_d^*(y_p) \subset [y_\ell, y_r]$, we know y_d^* is increasing on $[y_\ell, y_r]$.

[19] We further analyze ordering in the *competitive delegation* extension and show in Proposition 2 that y_d^* is increasing on X under standard assumptions.

Broadly, centrists want their ally, moderates want a more centrist moderate, and extremists want someone on their side of the median. Relative to the more general case, the characterization is sharper in two key ways. First, a centrist principal, $y_p = y_m$, always wants an ally. Any other representative proposes farther policy *and* also expands the acceptance set on both sides. Second, moderate principals never want a centrist representative. The appeal of a centrist requires that they shift the median, but that is not possible in this case.

Quadratic Utility in a Polarized Legislature

We now sharpen the general characterization by studying a special case in which all players have quadratic policy utility (Assumption 3) and almost all of the other politicians are extreme (Assumption 4).

Assumption 3. *For all $x, y \in X$, $u(x, y) = 1 - (x - y)^2$.*[20]

Assumption 4. *For all $y_d \in X$, $A(y_d) \cap \{y_1, y_2, ..., y_{k-1}, y_k\} = \{y_\ell, y_r\}$.*

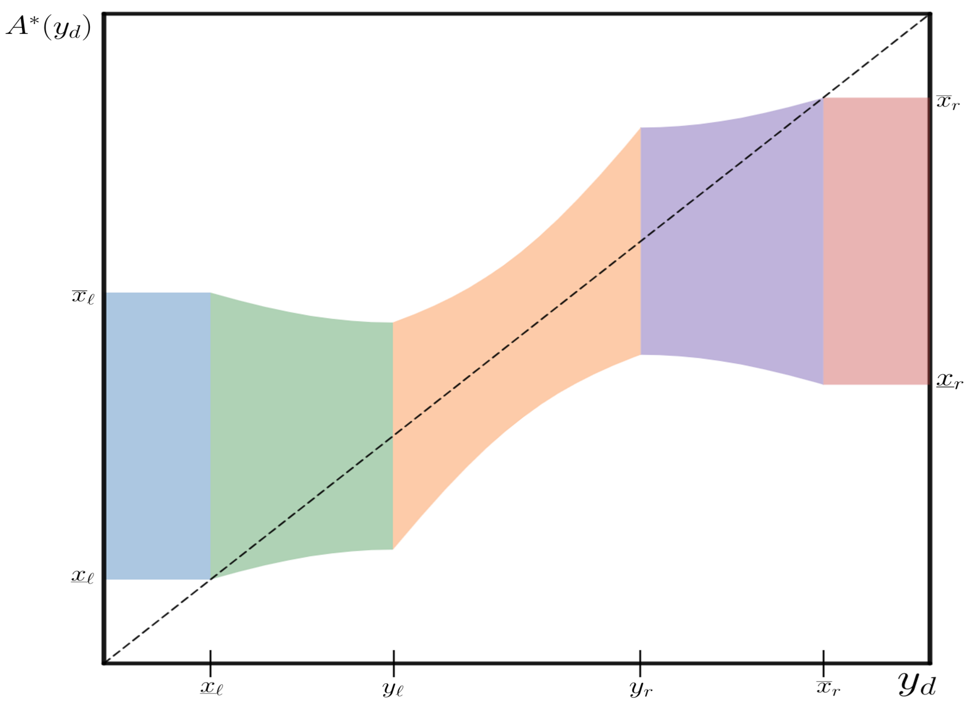
More precisely, Assumption 4 says (i) y_ℓ and y_r are centrist enough to be inside the acceptance set for all y_d, and (ii) all other legislators are extreme enough to always be outside. Essentially, it strengthens Assumptions 1 and 2, which guarantee at least one politician will always be constrained by the acceptance set. Now, in addition, the set of constrained proposers in K is constant in y_d. Let ρ_L denote aggregate proposal rights of left extremists, define ρ_R analogously, and let $\rho_E \equiv \rho_L + \rho_R$.

These assumptions yield a variety of additional insights. First, a wide interval of principals each have a unique optimal representative. Second, a unique centrist strictly prefers an ally. Third, a centrally-located interval wants centrists, flanked by two intervals who want moderates, and the rest want extremists. Fourth, there is a *dead zone* of representatives around y_ℓ or y_r who are not optimal for anyone. Finally, representatives are globally ordered.

[20]Since our policy space is $X = [0, 1]$, we include the constant 1 to satisfy the maintained *bad status quo* assumption that all policies provide greater utility than no agreement.

Crucially, Assumptions 3 and 4 allow us to strengthen Lemma 3 and more precisely characterize how the acceptance set varies with y_d.[21] Over extremists, we already know (i) $y_d \leq \underline{x}_\ell$ implies $A(y_d) = [\underline{x}_\ell, \overline{x}_\ell]$ and (ii) $y_d \geq \overline{x}_r$ implies $A(y_d) = [\underline{x}_r, \overline{x}_r]$ — where $\underline{x}_\ell < \underline{x}_r$ and $\overline{x}_\ell < \overline{x}_r$. Now, we also know $\underline{x}_r < \overline{x}_\ell$, since Assumption 4 implies that all centrists are always in the acceptance set, i.e., $[y_\ell, y_r] \subset A(y_d)$ for all y_d. Over each interval of moderates, $(\underline{x}_\ell, y_\ell)$ and $(y_r, \overline{x}_r)$, we already know $A(y_d)$ shrinks as y_d shifts inward and that effect vanishes as y_d approaches the nearest centrist. And over centrists, $[y_d, y_r]$, we already know that the acceptance set shifts monotonically with y_d. Now, we also know that on each interval $(\underline{x}_\ell, y_\ell)$, $(y_r, \overline{x}_r)$, and (y_ℓ, y_r), the upper bound is convex, the lower bound is concave, and both are continuously differentiable.[22] Figure 3 illustrates.

Figure 3: How the acceptance set varies with y_d



Note: Figure 3 illustrates how $A^*(y_d)$ varies under Assumptions 3 and 4. It is continuous and (i) constant for $y_d \leq \underline{x}_\ell$, (ii) contracting inward over $y_d \in (\underline{x}_\ell, y_\ell)$ at a decreasing rate that goes to zero as y_d approaches y_ℓ, (iii) increasing over (y_ℓ, y_r), (iv) expanding outward over $(y_r, \overline{x}_r)$ at an increasing rate that starts from zero near y_r, and (v) constant over $\overline{x}_r$. Additionally, $\underline{x}_\ell < \underline{x}_r < \overline{x}_\ell < \overline{x}_r$.

[21] See Lemma A3.

[22] At y_ℓ and y_r there are kinks in both bounds of the acceptance set because it shifts more sharply with y_d over centrists.

In turn, we can sharpen our characterization of P's objective function and her optimal representatives. Specifically, the convexity properties of the acceptance correspondence ensure that P's objective function is strictly quasi-concave on each interval of moderates and centrists, which implies that in each P has a unique locally optimal representative.

Proposition 2 sharpens our earlier characterization of optimal representatives y_d^*. First, each centrist has a unique optimal representative, who is also a centrist. Next, each moderate wants either a unique moderate — who is biased inwards — or a unique centrist in (y_ℓ, y_r). Their preference between these options is characterized by two cutpoints, one on each side of the spectrum. Those outside want their optimal moderate and those inside want their optimal centrist, with indifference at the cutpoint. Finally, each extremist wants either a unique moderate or any aligned extremist, with two cutpoints similarly partitioning their preference between these two options. Those outside prefer any aligned extremist but those inside prefer their optimal moderate. Figure 4 illustrates.

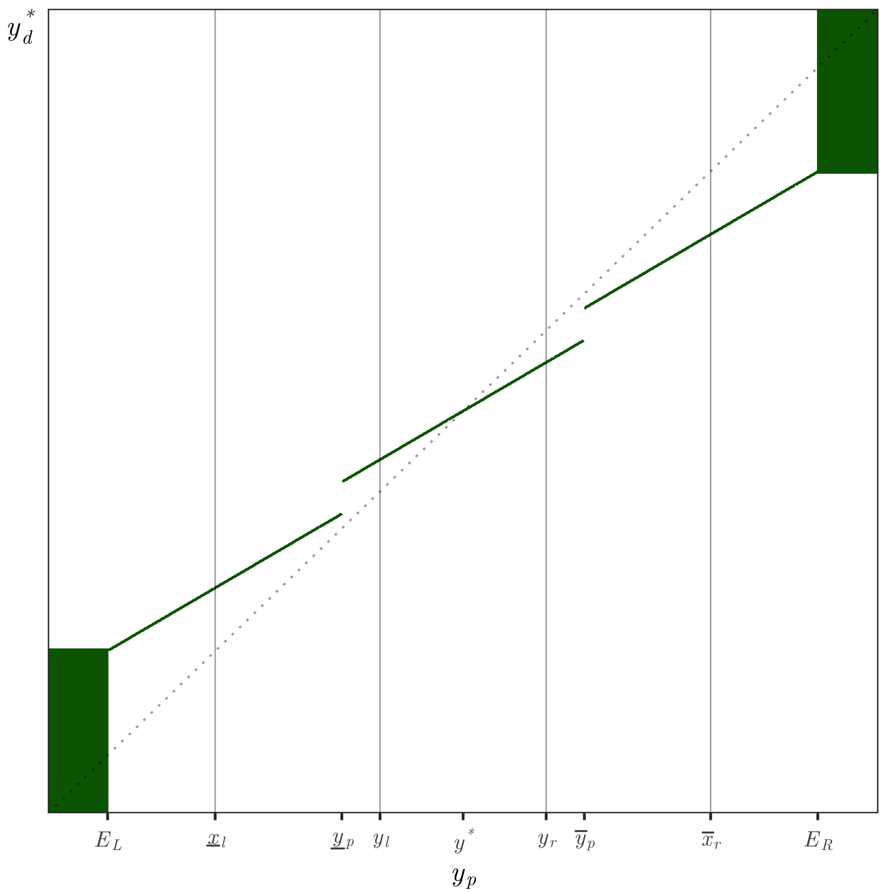
Proposition 2. *Under Assumptions 3 and 4, the correspondence y_d^* is increasing. Moreover, there are intervals (E_L, E_R) and $(\underline{y}_p, \overline{y}_p)$ satisfying $[y_\ell, y_r] \subset [\underline{y}_p, \overline{y}_p] \subset (\underline{x}_\ell, \overline{x}_r) \subset (E_L, E_R)$ such that y_d^* is single-valued and continuous on $(E_L, E_R)\backslash\{\underline{y}_p, \overline{y}_p\}$. Additionally,*

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

Furthermore, $y_p = \underline{y}_p < y_\ell$ implies $y_d^*(y_p)$ is a doubleton with $\min y_d^*(y_p) < y_\ell < \max y_d^*(y_p)$, and $y_p = \underline{y}_p = y_\ell$ implies $y_d^*(y_p) = y_\ell$; and symmetrically for $y_p = \overline{y}_p$.

Proposition 2 characterizes which principals want a centrist, moderate, or extremist representative. Those preferring a centrist are in a centrally-located interval, $(\underline{y}_p, \overline{y}_p)$, including

Figure 4: Optimal representatives



Note: Figure 4 illustrates key properties of the optimal representative characterization in Proposition 2: (i) very extreme principals, $y_p \notin (E_L, E_R)$, prefer any aligned extremist; (ii) intermediate principals, $y_p \in (E_L, \underline{y}_p) \cup (\overline{y}_p, E_R)$ prefer a unique moderate who is strictly more centrist; and (iii) centrist principals, $y_p \in (\underline{y}_p, \overline{y}_p)$, prefer a unique centrist.

some moderates, potentially on both sides. Those who prefer a moderate are on either side of that interval, in two intermediate intervals, $(E_L, \underline{y}_p)$ and $(\overline{y}_p, E_R)$, each including some extremists. Those who prefer an extremist are even further out, in two intervals, $[0, E_L)$ and $(E_R, 1]$, containing only extremists.[23] Finally, P's optimal representative is unique almost everywhere in (E_L, E_R), the potential exceptions being $\underline{y}_p$ or $\overline{y}_p$, where P is indifferent between a centrist and a moderate.[24]

[23]Either interval may be empty. For example, all left extremists may strictly prefer a moderate if ρ_R is sufficiently high (see Lemma A7 in the appendix).

[24]Notice also that y_d^* is also continuous a.e. on X. Lower hemi-continuity fails only if $y_p \in \{E_L, \underline{y}_p, \overline{y}_p, E_R\}$,

Proposition 2 also reveals there are representatives who are *not* optimal for *any* principal. This *dead zone* consists of open interval(s) around at least one of y_ℓ or y_r, possibly both. Essentially, these representatives are not optimal for any principal because they do not constrain powerful extremists enough relative to other representatives near them. On at least one side, principals near the moderate-centrist boundary want to bias inward in order to constrain their aligned extremist proposers. Then, the indifferent moderate principal wants either a strict centrist or strict moderate because representatives around that boundary do not constrain P's aligned extremists enough to justify their worse proposals.

Next, we show that an ally is uniquely optimal for exactly one principal: a centrist whom we denote y^*. Moreover, all principals in (E_L, E_R) want to bias their representative towards y^*, so we refer to it as the *locus of attraction.* Additionally, due to the monotonicity properties of y_d^*, they bias in an ordered way.

By Proposition 1, only centrists may strictly prefer an ally representative. If $y^* \in (y_\ell, y_r)$, then shifting y_d away from $y_p = y^*$ must not change P's expected payoff. In this case, y^* is the location where shifting y_d does not change the expected distance from y^* to the boundaries. Specifically, the marginal effect of y_d on P's value combines the direct effect through d's proposal with indirect effects through the boundary proposals. The direct effect is zero for all $y_p \in [y_\ell, y_r]$. In general, the indirect effects can be positive or negative — since both boundaries shift in the same direction, P will gain on one side but lose on the other. At $y_p = y^*$, however, these effects must also equal zero. Formally,

$$0 = \frac{\partial U(y_d, y_p)}{\partial y_d}\Big|_{y_d = y_p} = \rho_L \frac{\partial u(\underline{x}(y_d), y_p)}{\partial y_d}\Big|_{y_d = y_p} + \rho_R \frac{\partial u(\overline{x}(y_d), y_p)}{\partial y_d}\Big|_{y_d = y_p}. \tag{5}$$

Essentially, P's marginal gain from shifting one boundary towards y_p must exactly offset her marginal loss from shifting the other boundary away. Furthermore, since d is the median if $y_d = y^* = y_p$, both boundaries will be equally far from y_p so P is indifferent between them

where the optimal representative switches from a moderate to an extremist or centrist.

and (5) reduces to:

$$\rho_L \frac{\partial \underline{x}(y_d)}{\partial y_d}\Big|_{y_d=y^*} - \rho_R \frac{\partial \overline{x}(y_d)}{\partial y_d}\Big|_{y_d=y^*} = 0. \tag{6}$$

We define a function that represents the effect in (6) as a function of y_p and thus exactly coincides at $y_p = y^*$. Specifically, define $\lambda : [y_l, y_r] \to \mathbb{R}$ as

$$\lambda(y_p) \equiv \rho_L \frac{\partial \underline{x}(y_d)}{\partial y_d}\Big|_{y_d=y_p} - \rho_R \frac{\partial \underline{x}(y_d)}{\partial y_d}\Big|_{y_d=y_p} \tag{7}$$

for $y_p \in (y_\ell, y_r)$, then set $\lambda(y_\ell) = \lim_{y_p \to y_\ell^+} \lambda(y_p)$ and $\lambda(y_r) = \lim_{y_p \to y_r^-} \lambda(y_p)$. This function has two properties that together help us characterize y^*. First, it is strictly decreasing in y_p due to the concavity of $\underline{x}(y_d)$ and convexity of $\overline{x}(y_d)$. Second, its sign indicates which way P wants to bias y_d. For example, $\lambda(y_p) > 0$ implies rightward bias is optimal — since shifting y_d rightward from y_p will make P better off by decreasing the expected distance between y_p and boundary proposals.

Proposition 3 uses λ to show that y^* is unique and provide simple conditions to locate it.

Proposition 3. *Under Assumptions 3 and 4, $y_d^*|_{(E_L,E_R)}$ 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, $y^* \in (y_\ell, y_r)$.*

On both sides, fringe centrists want to bias inwards — implying $y^* \in (y_\ell, y_r)$ — if the signs of $\lambda(y_\ell)$ and $\lambda(y_r)$ differ. Otherwise, all centrists want to bias in the same direction, so y^* is on the boundary that does not want to bias inward.

We know from properties of λ that all centrists bias towards y^*. Moreover, all moderates bias inwards towards y^* and extremists never strictly prefer to bias outwards. Thus, P always biases towards y^*.

Corollary 3.1. *All $y_p \in (E_L, E_R)$ strictly prefer to bias their representative towards y_p^*.*

Moreover, principals closer to y^* want a representative who is closer to y^*. This property follows from monotonicity properties of y_d^*. Optimal representatives are monotonic over centrist principals in the general setting, but now they are globally monotonic.

Proposition 4 uses λ to characterize (i) whether any moderates on either side want a centrist representative and (ii) which side has the *dead zone*, denoted Δ.

Proposition 4. *Under Assumptions 3 and 4,*

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* $y_\ell, y_r \in \Delta$.

The key factor underlying Proposition 4 is whether centrists at y_ℓ or y_r want to bias inwards. If either does, then its nearby moderates also want to bias inward enough to have a centrist representative. Thus, the sign of $\lambda(y_\ell)$ characterizes whether $\underline{y}_p < y_\ell$ and similarly $\lambda(y_r)$ characterizes whether $y_r < \overline{y}_p$. If $\lambda(y_r) < 0 < \lambda(y_\ell)$, then moderates near y_ℓ want to bias rightward into (y_ℓ, y_r) and symmetrically for moderates near y_r, so $y^* \in (y_\ell, y_r)$. If not, then one of y_ℓ or y_r wants an ally representative — i.e., $y^* \in \{y_\ell, y_r\}$ — and none of their nearby moderates want a centrist, but some moderates on the other side will want a centrist.

Proposition 4 implies that representatives at y_ℓ and y_r can be optimal for (i) nobody, (ii) exactly one P, or (iii) an interval of centrist P. Notably, they are the only representatives who can be uniquely optimal for more than one P.

We can further sharpen our characterization of y_d^* if the median is fixed ($y_\ell = y_r$) and the proposal power of extremists is balanced ($\rho_L = \rho_R$). Under these conditions, y_d^* is continuous and over (E_L, E_R) it is a convex combination of y_p and y_m, where $\delta\rho_E$ is the weight on y_m.

Corollary 4.1. *Suppose* $y_\ell = y_r$ *and* $\rho_L = \rho_R$*. Assumptions 3 and 4 imply* $y_m = y^* = \underline{y}_p = \overline{y}_p$ *and* $y_d^*(y_p)\big|_{(E_L,E_R)} = (1 - \delta\rho_E)y_p + \delta\rho_E y_m$.

Using Corollary 4.1, we can shed light on how far P moderates, $|y_d^*(y_p) - y_p|$. The effect of ρ_E highlights the key force for moderation: P moderates to constrain extremists and thus moderates more as extremists gain proposal power. Additionally, more centrist principals do not moderate as far, since biasing y_d towards y_m has a weaker effect on $A^*(y_d)$. Essentially, the "price" of moderating extremist proposals rises as y_d gets closer to y_m. Finally, increasing δ induces P to moderate further. Essentially, there is a drop in the "price" of constraining extremists — increasing δ makes m's expectations about future policymaking more prominent in her voting calculus and thus magnifies the effect of y_d on the acceptance set.

Effects of Extremism. We have shown an incentive to use strategic representation to counteract extremists. We now show how that varies with changes in relative extremism. Specifically, Proposition 5 characterizes how shifting proposal rights between L and R affects: the locus of attraction, y^*; the set principals who choose a moderate, (E_l, E_R); and the set of principals who choose a centrist, $(\underline{y}_p, \overline{y}_p)$. Figure 5 illustrates.

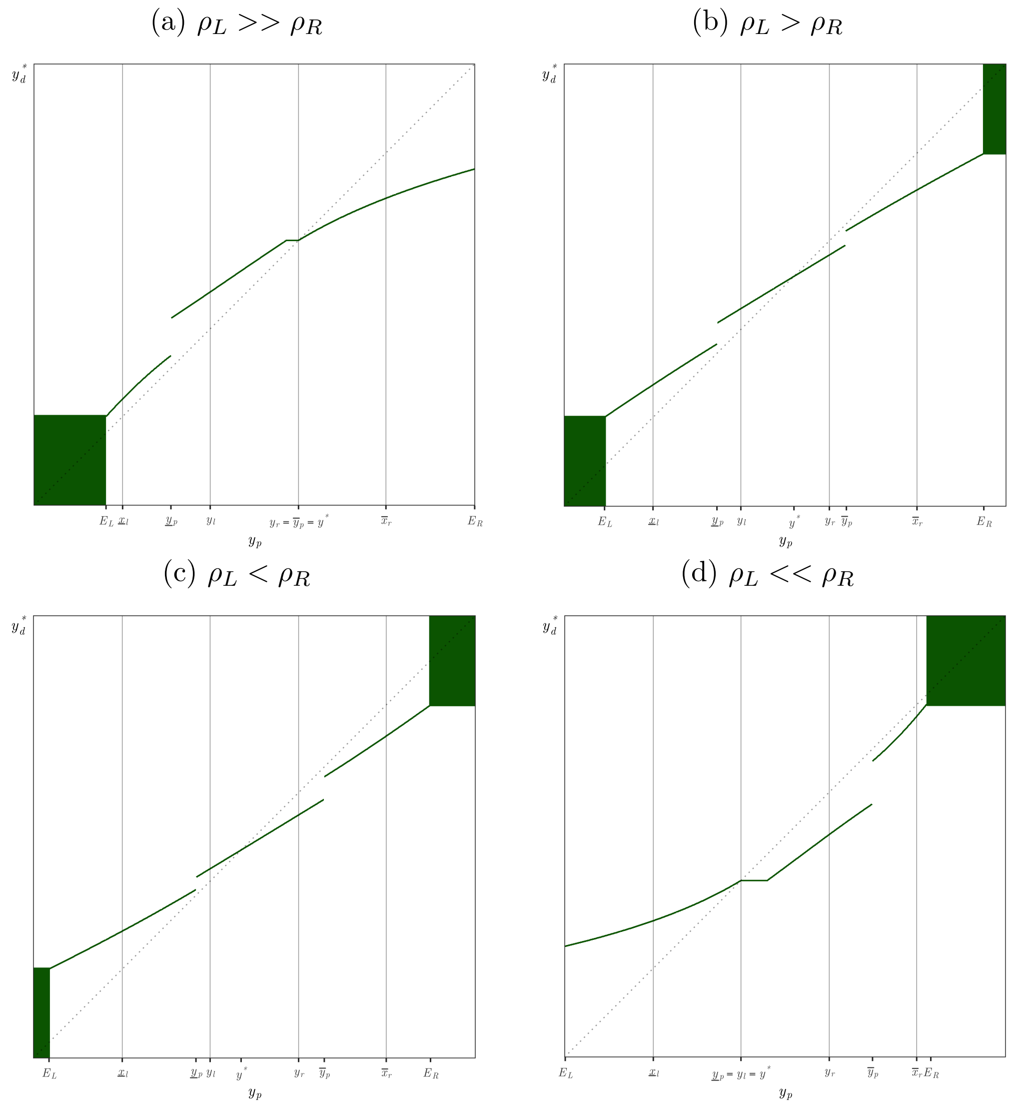
Proposition 5. *Fix $\rho_E = \rho_L + \rho_R$. Under Assumptions 3 and 4, increasing ρ_L:*

1. *weakly increases the locus of attraction, y_p^*;*
2. *weakly increases the set of principals who strictly prefer a non-extremist, (E_L, E_R); and*
3. *weakly decreases the set of principals who strictly prefer a veto player, $(\underline{y}_p, \overline{y}_p)$.*

Given ρ_E, transferring recognition probability between L and R does not affect the acceptance set on its own, since the median is indifferent between their proposals. Yet, this transfer does affect P's delegation incentives. For example, increasing ρ_L at ρ_R's expense amplifies P's sensitivity to constraining left extremists but also dampens her sensitivity to constraining right extremists, and vice versa. Depending the location of y_p, these effects can change P's optimal representative in different ways.

First, y_p^* shifts rightward since increasing ρ_L strengthens centrists' desire to constrain left extremists by skewing rightward. Essentially, centrists want to constrain both extremists, but grow more concerned about constraining the strengthened side and less concerned about the weakened side.

Figure 5: Optimal representatives vary with extremist proposal rights



Note: Figure 5 illustrates Proposition 5 to show how optimal representatives change as recognition probability is transferred between the extremist politicians.

Next, $\overline{y}_p$ and $\underline{y}_p$ both shift leftward since increasing ρ_L makes left moderates more inclined towards centrist representatives and right moderates less inclined towards them. Although moderates experience a similar effect on their strategic calculus as centrists, they weigh whether to choose an aligned moderate or a centrist. For moderates on the weakened side, biasing towards centrists is more appealing than before due to higher return from constraining their opposing extremist, but each centrist is less appealing due to higher cost of relaxing the far extremist's constraint. On the strengthened side, the forces are analogous but push in the other directions.

Finally, E_L and E_R both shift rightward since increasing ρ_L makes right extremists more inclined to moderate and left extremists less inclined. Extremists want to constrain their opposing extremist but not their aligned extremist. Thus, their desire to moderate varies with relative extremism differently if they are on the strengthened side rather than the weakened one. On the strengthened side, moderating is more appealing due to greater return from constraining their opposing extremist and also lower cost of constraining their aligned extremist. On the weakened side, moderating is less appealing since these effects reverse.

Extensions

Thus far, our analysis has focused on addressing the question of "who is best for one principal filling one position in a collective body?" To clarify and expand the scope of our results, we now extend the baseline setting in three ways to consider variants of that question. The first is *collective representation* — addressing "who would a group choose to fill the position?" The second is the *value of representation* —– addressing "how much does the principal gain from optimal representation?" Third, we study *competitive representation* —– addressing "what if multiple principals fill multiple positions?"

Collective Representation

Representatives are often chosen by groups —– e.g., voters, parties, etc. —– so we want to understand collective choice over representatives. We show when this collective choice coincides with choice by a single principal. Moreover, we show that a weak condition on extremist proposal rights ensures that coalitions have a natural ordering. Overall, our analysis has implications for electoral competition and voting into collective policymaking bodies.

The key step is verifying when collective choice always coincides with the choice of a single, *decisive* principal. Our most general setting does not guarantee this property.

Choosing between two representatives is a choice between policy lotteries, so preferences over representatives are not necessarily ordered in a way that yields a decisive principal. But with quadratic policy utility, those preferences are *order restricted* — i.e., for any pair of candidates y_d to y_d', the set of y_p for which P prefers y_d is an interval (Cho and Duggan 2003; Kartik et al. 2022). This property implies that, for example, under majority rule the median principal is decisive (Cho and Duggan 2003). More generally, a decisive principal exists under any *strong voting rule* (Duggan 2014), so collective choice over representatives reduces to individual choice under broad class of voting rules — e.g., the median under majority rule.

Additionally, we provide conditions ensuring that any pair of candidate representatives induces a cutpoint with all rightward principals always preferring the right candidate, and vice versa. Specifically, Proposition 6 shows that extremist proposal rights below a common threshold ensure that preferences satisfy the single-crossing condition (Milgrom and Shannon 1994) and thus coalitions are consistently ordered.

Proposition 6. *Under Assumptions 3 and 4, the principal's equilibrium value, U, satisfies the single-crossing condition if* $\max\{\rho_L, \rho_R\} \leq \frac{1}{2\delta}$.

Underlying Proposition 6 is that if their aligned extremists are very likely to propose, then extremist principals may choose the more extreme candidate over the closer candidate in some pairwise comparisons. Otherwise, they always prefer the closer candidate.

The taste for extremism can emerge only when comparing moderates. When choosing between two moderates from the same side of the distribution, extremists from the opposite side face a trade off. The more allied moderate candidate makes a more favorable proposal and constrains opposing extremists more than the alternative but also constrains aligned extremists. Due to risk aversion, the benefit of constraining opposing extremists outweighs the cost of constraining aligned extremists unless their aligned extremists have sufficiently high recognition probability. In contrast, when choosing between two centrist candidates, extremists prefer the nearest candidate because both extreme proposals move in the same direction as y_d within (y_ℓ, y_r).

Value of Representation

Next, we study how much the principal gains from having their optimal representative rather than an *ally* with the same ideal point.

Specifically, we define P's *value of representation* as

$$\nu(y_p) \equiv U(y_d^*(y_p), y_p) - U(y_p, y_p). \tag{8}$$

To sharpen our analysis, we focus on the same conditions as in Corollary 4.1: the polarized quadratic setting (Assumptions 3–4), a fixed median ($y_\ell = y_r$), and balanced extremist proposal rights ($\rho_L = \rho_R$).

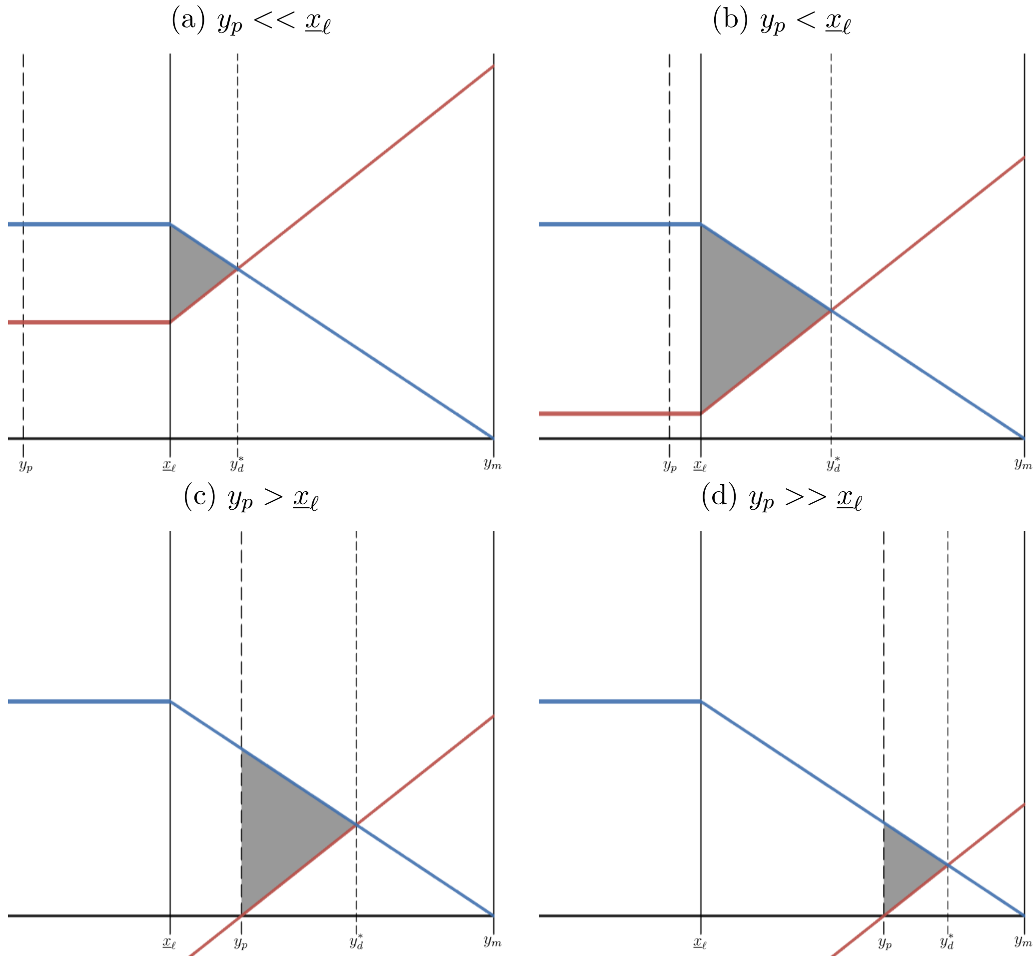
We show that extremists benefit more from optimal representation as they become more moderate and, conversely, moderates benefit more as they become more extreme. Thus, on each side of the spectrum, the value of representation is highest for principals on the extremist-moderate boundary — i.e., $y_p \in \{\underline{x}_\ell, \overline{x}_r\}$. Figure 6 illustrates.

Proposition 7. *Under Assumptions 3 and 4, if $\rho_\ell = \rho_r$ and $\rho_L = \rho_R$, then $\nu(y_p)$ is strictly increasing on $[E_L, \underline{x}_\ell]$, strictly decreasing on $[\underline{x}_\ell, y_m]$ and analogously for $y_p \in [y_m, E_R]$.*

To characterize ν, we exploit the property that P's optimal representative, $y_d^*(y_p)$, always balances her marginal benefit of moderation against her marginal cost. To illustrate more precisely, if a left-leaning P moderates further from any $y_d \in (\underline{x}_\ell, y_m)$ she enjoys marginal benefit $\frac{\delta\rho_E\rho_d(y_m - y_d)}{1-\delta\rho_E}$, which is her gain from further constraining extremist proposals, and incurs marginal cost $\rho_d(y_d - y_p)$, which is her loss from shifting d's proposal further away. Since (i) P's marginal benefit decreases in y_d over this interval and is independent of y_p, while (ii) P's marginal cost increases in y_d and decreases in y_p,[25] her marginal benefit exceeds marginal cost for $y_d < y_d^*(y_p)$ and vice versa, with the difference increasing in their distance. Thus, $\nu(y_p)$

[25] As y_p moderates, the marginal benefit from constraining R's proposal decreases relative to constraining L. Since the boundaries of A^* contract at the same rate, if $\rho_L = \rho_R$ then the loss in marginal benefit from shifting $\overline{x}(y_d)$ is exactly equal to the gain from shifting $\underline{x}(y_d)$.

Figure 6: How the value of representation varies with y_p



Note: Figure 6 displays the value of representation (ν) for four different values of $y_p \in (E_L, y_m)$. In each panel, $\nu(y_p)$ equals the area of the shaded region between the two curves, which are marginal benefit (downward sloping) and marginal cost (upward sloping) of moderation as functions of y_d.

equals the area between P's marginal cost and benefit curves over $y_d \in [\max\{\underline{x}_\ell, y_p\}, y_d^*(y_p)]$.[26]

If P is sufficiently extreme, $y_p \notin (E_L, E_R)$, she nominates an aligned extremist and thus $\nu(y_p) = 0$. As P becomes less extreme, (i) her marginal cost curve shifts down and (ii) y_d^* shifts towards y_m, which increases the difference between marginal benefit and marginal cost

[26]This follows from the fundamental theorem of calculus:

$$\nu(y_p)\big|_{(E_L,\underline{x}_\ell)} = \int_{\underline{x}_\ell}^{y_d^*(y_p)} \left(\frac{\delta \rho_E \rho_d (y_m - y_d)}{1 - \delta \rho_E} - \rho_d (y_d - y_p) \right) \mathrm{d}y_d,$$

and

$$\nu(y_p)\big|_{(\underline{x}_\ell, y_m)} = \int_{y_p}^{y_d^*(y_p)} \left(\frac{\delta \rho_E \rho_d (y_m - y_d)}{1 - \delta \rho_E} - \rho_d (y_d - y_p) \right) \mathrm{d}y_d.$$

at all $y_d \in [\underline{x}_\ell, y_d^*(y_p)]$, so her value of delegation rises. Figure 6a–6b illustrates.

For moderate P, the acceptance set induced by their ally will shrink as y_p approaches y_m, so there is a smaller difference between P's marginal benefit and marginal cost at every $y_d \geq y_p$. Since the extent of P's optimal bias also decreases, the value of representation decreases as y_p approaches y_m. Figure 6c–6d illustrates.

Competitive Representation

Thus far, we have focused on a principal filling one position and fixed the rest of the political environment. This can reflect situations in which other politicians are already in office, but our analysis also highlights that incentives for moderation will arise in situations where multiple positions will be filled simultaneously (as noted by, e.g., Gailmard and Hammond 2011). In this section, we explore whether those incentives will strengthen or weaken by extending our baseline setup so that two principals simultaneously pick their representatives.

We extend the model to have two principals, P_a and P_b, each simultaneously appointing representatives, a and b, to fill two positions in a five-player body. The three other politicians are two extremists, L and R, and a veto player, M, who determines whether any proposal passes. Finally, we maintain Assumptions 3 and 4 throughout and assume $y_L < y_{p_a} < y_M < y_{p_b} < y_R$, where y_{p_a} and y_{p_b} denote the principals' ideal points.

By Lemma 1, each $y_a, y_b \in X^2$ induces a unique distribution over policy outcomes characterized by the equilibrium acceptance set, $A^*(y_a, y_b)$. To streamline key points, we assume (i) y_{p_a} and y_{p_b} are both always inside the acceptance set, while (ii) y_L and y_R are always outside.

Our characterization of optimal representatives in the baseline analysis also characterizes best responses in this competitive setting. Since both principals are moderates, each will bias their representative towards M in equilibrium, so $y_{p_a} < y_a^* < y_M < y_b^* < y_{p_b}$. Furthermore, with quadratic policy utility, Proposition 2 implies that each principal always has a unique best response. Specifically, principal P_a's best response to y_b, denoted $y_a(y_b)$, is the unique

$y_a \in (y_{p_a}, y_M)$ satisfying the first-order condition:

$$\frac{\partial \underline{x}(y_a, y_b)}{\partial y_a}[\rho_L(y_{p_a} - \underline{x}(y_a, y_b)) + \rho_R(\overline{x}(y_a, y_b) - y_{p_a})] - \rho_a(y_a - y_{p_a}) = 0, \quad (9)$$

and P_b's best response function is analogous.

Lemma 5 establishes that each principal's best response is monotone. Moreover, the direction is determined by which extremist has greater proposal rights

Lemma 5. *If $\rho_L < \rho_R$, then y_a is strictly decreasing and y_b is strictly increasing; and vice versa if $\rho_L > \rho_R$. 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}.*

Lemma 5 implies the principals' best responses intersect once. Thus, a unique pair of representatives is mutually optimal and each is strictly more centrist than their principal.

Proposition 8. *There is a unique equilibrium, in which $y_a^* \in (y_{p_a}, y_m)$ and $y_b^* \in (y_m, y_{p_b})$.*

Additionally, Lemma 5 implies that the principal aligned with weaker extremists will moderate further in the competitive setting than in the baseline setting, whereas the principal aligned with stronger extremists will moderate less.

Corollary 8.1. *In equilibrium: (i) $\rho_L < \rho_R$ implies $y_a(y_b) < y_a^* < y_M < y_b(y_a) < y_b^*$; (ii) $\rho_L = \rho_R$ implies $y_a(y_b) = y_a^* < y_M < y_b(y_a) = y_b^*$; and (iii) $\rho_L > \rho_R$ implies $y_a^* < y_a(y_b) < y_M < y_b^* < y_b(y_a)$.*

Corollary 8.1 is driven by the two effects of opponent moderation. To fix ideas, consider shifting y_b inwards. One effect is that extremist proposals also shift inwards, which directly benefits P_a and decreases her marginal benefit from shifting y_a inward. Through this effect, moderation by P_b *substitutes* for moderation by P_a. The other effect is that the acceptance set becomes more sensitive to y_a, i.e., $\frac{\partial^2 \underline{x}(y_a, y_b)}{\partial y_a \partial y_b} < 0$. Through this channel, moderation by P_b reduces the "price" of moderating extremist proposals and *complements* moderation by P_a.

Which effect dominates depends on the balance of extremist proposal rights. The complementary effect dominates on the weak side and conversely on the strong side. If

$\rho_L = \rho_R$, then the increased marginal elasticity of A^* to y_{d_i} exactly offsets the decreased marginal benefit of moderation. As one extremist gains proposal power at the expense of the other, since each principal's aligned extremist is closer to her ideal point than the non-aligned extremist, the marginal benefit of moderation declines in $y_{d_{-i}}$ at a slower rate for the weak-side principal than the strong-side principal.

Conclusion

We study representatives who participate in collective policymaking. A key force in our analysis is that a representative's ideology affects legislature-wide expectations about policymaking. This force is present in many contexts and we study its consequences for representation. We show how it has important *anticipation effects* by shaping exactly which policies each politician will support, thereby influencing what would pass and what extremists will propose.

We provide a general logic for why moderate representatives can be appealing. We show that (i) all centrist principals want a centrist representative who will be the median (de facto veto) politician and (ii) all moderate principals want a more centrist representative. Even when they are not the median, their closer alignment improves the median's expectation about proposals and thus narrows what can pass, which constrains extremist politicians.

Moreover, under standard assumptions, principals who are not too extreme have a uniquely optimal representative, who is biased inward towards a centrally located *locus of attraction*. Additionally, we find a *dead zone* of representatives on the centrist-moderate fringe(s) who are not optimal for any principal. Furthermore, the locations of the dead zone and locus of attraction depend on the balance of extremist proposal rights. As that balance changes, the principal grows more concerned with constraining extremists on the gaining side — so the the dead zone grows on that side while the locus of attraction shifts away.

We focused on a collective policymaking environment governed by simple majority rule. Our results generalize to any *strong* voting rule, since there will be a single decisive principal who will effectively determine what can pass (Duggan 2014). To illustrate, our analysis with a fixed median ($y_\ell = y_r$) in Corollary 1.2 is equivalent to the principal appointing a proposer into a dictatorial rule setting with the dictator already in place. Future work could study representation in settings where more than one politician will be decisive. For example, under supermajority rules the acceptance set is determined by two (endogenous) veto players rather than a single median. In those settings, extending our analysis requires characterizing an implicit curve defined by a system of nonlinear equations.

Our main results shed new light on representation in separation-of-powers systems and congressional committees. Additionally, our extensions provide insight into mass representation, with implications for studying behavior by voters and elites in elections for positions in collective bodies. Future work in this direction could build on our foundations in order to explore how elite polarization and extremism affect elections.

Appendix

Contents

A General Setting

A.1 Proof of Lemma 1.

Follows from Propositions 1–2 in Cardona and Ponsati (2011). □

A.2 Define the cutpoints $\underline{x}_\ell, \overline{x}_\ell, \underline{x}_r$, and $\overline{x}_r$.

From Cardona and Ponsati (2011), the equation

$$u(x, y_\ell)(1 - \delta\rho_d) - \delta \left(\sum_{i \in K : |y_i - y_\ell| \geq |y_\ell - x|} \rho_i u(x, y_\ell) + \sum_{i \in K : |y_i - y_\ell| < |y_\ell - x|} \rho_i u(y_i, y_\ell) \right) = 0 \qquad (10)$$

has exactly two solutions: $\underline{x}_\ell \in (0, y_\ell)$ and $\overline{x}_\ell \in (0, y_\ell)$. Similarly, $\underline{x}_r \in (0, y_r)$ and $\overline{x}_r \in (0, y_r)$ are the only solutions of

$$u(x, y_r)(1 - \delta\rho_d) - \delta \left(\sum_{i \in K : |y_i - y_r| \geq |y_r - x|} \rho_i u(x, y_r) + \sum_{i \in K : |y_i - y_r| < |y_r - x|} \rho_i u(y_i, y_r) \right) = 0. \qquad (11)$$

A.3 Proof of Lemma 2.

Part 1 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 (10). Thus, $\underline{x}(y_d) = \underline{x}_\ell$. Analogously using (11), $y_d \geq \max A^*(y_d) = \overline{x}(y_d)$ 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 $y_d \leq \overline{x}_\ell$. Then, uniqueness of $A^*(y_d)$ implies that $y_d \notin \text{int } A^*(y_d)$ is equivalent to the lower solution of (10) satisfying $\underline{x}_\ell \geq y_d$. Thus, $y_d \leq \overline{x}_\ell$ 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 $y_d \in (\underline{x}_\ell, \overline{x}_r)$. □

A.4 Proof of Lemma 3.

We first establish a useful property in Lemma A1.

Lemma A1. *$\underline{x}_\ell \le \underline{x}_r$ and $\overline{x}_\ell \le \overline{x}_r$ with both inequalities strict if $y_\ell < y_r$.*

Proof. To prove Lemma A1, we construct a function $\zeta(x,y) : [0,1] \times [y_\ell, y_r] \to \mathbb{R}$ which for each $y \in [y_\ell, y_r]$, has two unique roots, $\underline{x}(y)$ and $\overline{x}(y) = 2y - \underline{x}(y)$. We then establish that $\underline{x}(y)$ is continuous in y with $\underline{x}(y_j) = \underline{x}_j$ for $j \in \{\ell, r\}$. Having characterized $[\underline{x}_\ell, \overline{x}_\ell]$ and $[\underline{x}_r, \overline{x}_r]$ in this way, we prove Lemma A1 by showing that $\underline{x}(y)$ and $\overline{x}(y)$ are strictly increasing.

Let

$$\zeta(x,y) := u(x,y) - \delta\left(\sum_{i\in K:|y-y_i|<|y-x|} \rho_i u(y_i,y) + \left(\rho_d + \sum_{i\in K:|y-y_i|\ge|y-x|} \rho_i\right) u(x,y)\right).$$

Then, (i) $y_d \le y_\ell$ implies that $\zeta(x, y_\ell) = 0$ is equivalent to (10) and (ii) $y_d \ge y_r$ implies that $\zeta(x, y_r) = 0$ is equivalent to (11). Therefore $\zeta(\underline{x}_\ell, y_\ell) = 0$ and $\zeta(\overline{x}_r, y_r) = 0$. Trivially, $y_\ell = y_r$ implies $[\underline{x}_\ell, \overline{x}_\ell] = [\underline{x}_r, \overline{x}_r]$. To establish the strict inequalities for $y_\ell < y_r$, we use the following properties of $\zeta(x,y)$: (i) it is continuous in each argument; (ii) it is differentiable in each argument at all (x,y) s.t. $|y_i - y| \ne |y - x|$ for any $i \in K$; (iii) strictly increasing in x and decreasing in y if $x < y$, strictly decreasing in x and increasing in y if $x > y$; (iv) $\zeta(x,y) > 0$ if $x = y$; and (v) $\zeta(0,y), \zeta(1,y) < 0$ (under Assumption 1).

It follows that for every $y \in [y_\ell, y_r]$, unique $\underline{x}(y) \in (0,y)$ and $\overline{x}(y) = 2y - \overline{x}(y) \in (y,1)$ exist such that $\zeta(\underline{x}(y), y) = \zeta(\overline{x}(y), y) = 0$. Moreover, $\underline{x}(y)$ and $\underline{x}(y)$ are continuous and differentiable almost everywhere.[27] At any y such that $\underline{x}(y)$ and $\overline{x}(y)$ are differentiable,

$$\underline{x}'(y) = 1 - \left[\delta \sum_{i\in N:|y-y_i|<|y-\underline{x}|} \rho_i \frac{\partial u(y_i,y)}{\partial y}\right]\left[\left(1 - \delta(\rho_d + \sum_{i\in N:|y-y_i|\ge|y-\underline{x}|} \rho_i)\right)\frac{\partial u(\underline{x},y)}{\partial y}\right]^{-1}.$$

[27]Continuity follows from the continuity of ζ. For differentiability a.e., notice that on every non-empty subset of $[y_\ell, y_r]$, symmetry requires that at most one root is constant in y. Now suppose an interval $(a,b) \subset [y_\ell, y_r]$ exists such that $\underline{x}(y)$ is constant. For $y, y' \in (a,b)$ where $y \ge y'$, it follows that $\overline{x}(y') - \overline{x}(y) = 2(y' - y) > 0$. But then $\zeta(\underline{x}(y'), y') < \zeta(\underline{x}(y), y) = 0$. Thus no interval exists where $\underline{x}(y)$ is constant in y. An analogous result holds for $\overline{x}(y)$. Therefore the set of $y \in [y_\ell, y_r]$ such that $y_i \in \{\underline{x}(y), \overline{x}(y)\}$ for some $i \in K$ is finite. The antecedent conditions of the implicit function theorem hold almost everywhere.

The strict concavity and symmetry of $u(x,y)$ imply that $\left|\frac{\partial u(x,y)}{\partial y}\right| > \left|\frac{\partial u(y_i,y)}{\partial y}\right|$ for all $y_i \in K$ such that $|y_i - y| < |x-y|$. Therefore $\underline{x}'(y) \in (0,2)$ and $\overline{x}'(y) = 2 - \underline{x}'(y) \in (0,2)$. The continuity of ζ therefore implies that $\underline{x}(y)$ and $\overline{x}(y)$ are strictly increasing. Thus $\underline{x}_\ell < \underline{x}_r$ and $\overline{x}_\ell < \overline{x}_r$. $\square$

Proof of Lemma 3.

Lemma 2 implies Parts 1 and 5. Part 3 is implied by Lemma 5 of Cardona and Ponsati (2011) which shows that $\underline{x}(y_d)$ and $\overline{x}(y_d)$ are strictly increasing in y_m if $A^*(y_d) \subset \text{int}X$. Parts 2 and 4 remain to be shown. We provide a proof for Part 2. The proof for Part 4 is analogous.

To prove Part 2, we first show that $\underline{x}_\ell(y_d)$ strictly increases and $\overline{x}_\ell(y_d)$ strictly decreases on $(\underline{x}_\ell, y_\ell)$. By Lemma 2, $y_d \in (\underline{x}_\ell, y_\ell)$ implies $\underline{x}(y_d)$ is the unique $x \in (\underline{x}_\ell, y_d)$ satisfying

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

and $\overline{x}(y_d) = 2y_\ell - \underline{x}(y_d)$. Notice that κ_ℓ is strictly increasing over $y_d \geq x$ and strictly decreasing over $y_d \in (x, y_\ell)$. Therefore $\underline{x}(y_d)$ is strictly increasing and $\overline{x}(y_d)$ strictly decreasing.

We now show that the rate of change of $\underline{x}_\ell(y_d)$ approaches zero as $y_d \to y_\ell$. Note that the set of x where $\left.\frac{\partial \kappa_\ell(x;y_d)}{\partial x}\right|_{x^-} \neq \left.\frac{\partial \kappa_\ell(x;y_d)}{\partial x}\right|_{x^+}$ is the finite set of x such that $x = y_i$ for some $i \in K$. It follows that $\underline{x}_\ell(y_d)$ is differentiable almost everywhere on $(\underline{x}_\ell, y_\ell)$. Therefore a non-empty interval $(y_\ell - \varepsilon, y_\ell)$ exists on which $\underline{x}(y_d)$ is continuously differentiable. At any y_d such that $\underline{x}_\ell(y_d)$ is differentiable,

$$\frac{\partial \underline{x}_\ell(y_d)}{\partial y_d} = \left(\frac{\delta\rho_d}{1-\delta[1-P(\overline{x})+P(\underline{x})]}\right)\left(\frac{\partial u(y_d,y_\ell)}{\partial y_d}\right)\left(\frac{\partial u(\underline{x},y_\ell)}{\partial \underline{x}}\right)^{-1},$$

where $\left(\frac{\delta\rho_d}{1-\delta[1-P(\overline{x})+P(\underline{x})]}\right)$ is a positive constant and $\underline{x}_\ell(y_d) < y_\ell$ implies $\frac{\partial u(\underline{x},y_\ell)}{\partial \underline{x}} > \frac{\partial u(y_d,y_\ell)}{\partial y_d}$. Therefore

$$\lim_{y_d\to y_\ell^-}\frac{\partial \underline{x}(y_d)}{\partial y_d} \propto \lim_{y_d\to y_\ell^-}\frac{\partial u(y_d,y_\ell)}{\partial y_d} = \left.\frac{\partial u(y_d,y_\ell)}{\partial y_d}\right|_{y_d=y_\ell} = 0.$$

Finally, Lemma A1 and Parts 1–5 imply that $A^*(y_d) \subset [\underline{x}_\ell, \overline{x}_r]$ for all y_d. $\square$

A.5 Proof of Lemma 4 and Proposition 1.

We establish general properties of $y_d^*(y_p)$ in Lemma A2, which yields Lemma 4 and Proposition 1 as corollaries.

Lemma A2. *Under Assumptions 1 and 2,*

1. $[0, \underline{x}_\ell] \cap y_d^*(y_p) \neq \emptyset \iff [0, \underline{x}_\ell] \in y_d^*(y_p)$ *and* $[\overline{x}_r, 1] \cap y_d^*(y_p) \neq \emptyset \iff [\overline{x}_r, 1] \in y_d^*(y_p)$*;*

2. $y_p \leq y_r$ *implies* $y_d^*(y_p) \cap (y_r, \overline{x}_r] = \emptyset$*, and* $y_p \geq y_\ell$ *implies* $y_d^*(y_p) \cap [\underline{x}_\ell, y_\ell) = \emptyset$*;*

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* $[y_p, \overline{x}_r] \cap y_d^*(y_p) = \emptyset$*;*

5. $y_p \in [0, \underline{x}(y_\ell)] \cup [\overline{x}(y_r), 1]$ *implies* $[y_\ell, y_r] \cap y_d^*(y_p) = \emptyset$*.*

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

$$U(y_d; y_p) := P(\underline{x}(y_d))u(\underline{x}(y_d), y_p) + [1 - P(\overline{x}(y_d))]u(\overline{x}(y_d), y_r) + \sum_{i \in N: y_i \in (\underline{x}(y_d), \overline{x}(y_d)]} \rho_i u(y_i, y_p).$$

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 $\underline{x}(y_d) \geq y_p$ (otherwise, $U(y_d - \varepsilon; y_p) > U(y_d; y_p)$ for some $\varepsilon > 0$). Next, by Lemmas A1 and 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 $y_p' \in y_d^*(y_p)$, a contradiction.

3. Lemma 3 implies that for $y_p < y_\ell$ we know that $U(y_d; y_p)$ is continuously differentiable on $(y_\ell - \varepsilon, y_\ell)$ with

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

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 increasing. 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, $\frac{\partial u(y_d; y_p)}{\partial y_p} > 0$ if $y_p < y_p$ and $\frac{\partial u(y_d; y_p)}{\partial y_p}\Big|_{y_d = y_p} = 0$. By continuity of $\frac{\partial u(y_d; y_p)}{\partial y_p}$ 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 parts 2 and 3, $y_\ell \in y_d^*(y_p)$ only if $y_p \geq y_r$.

□

B Quadratic Setting

B.1 Proof of Proposition 2.

First, we establish key properties in Lemmas A3–A7. Lemma A3 characterizes $A^*(y_d)$. Lemma A4 provides a sufficient condition for $U(y_d, y_p)$ to satisfy the single-crossing condition on $S \times X$, where $S \subseteq X$ is an arbitrary interval. Lemma A5 characterizes a compact interval in X for which that condition is satisfied. We use this result in our proofs of Propositions 2 and 6 by showing that $y_p^*(y_p)$ is a subset of this interval and must therefore be increasing. Lemmas A6 and A7 characterize local maxima of $U(y_d, y_p)$ for $y_d \in [\underline{x}_\ell, y_\ell]$ and $y_d \in [y_\ell, y_r]$, respectively. An analogous result for $y_d \in [y_r, \overline{x}_r]$ is omitted. We use these local maxima to characterize $y_d^*(y_p)$ in Proposition 2.

Lemma A3. *Under Assumptions 3–4:*

$$\underline{x}_\ell = y_\ell - \sqrt{\frac{1-\delta+\delta\rho_r(y_\ell-y_r)^2}{1-\delta(\rho_E+\rho_d)}}, \text{ and} \tag{12}$$

$$\overline{x}_r = y_r + \sqrt{\frac{1-\delta+\delta\rho_\ell(y_\ell-y_r)^2}{1-\delta(\rho_E+\rho_d)}}. \tag{13}$$

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

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

First, (12) follows from solving $y_d = \underline{x}(y_d)$ for $y_d < y_\ell$ and similarly (13) follows from solving $y_d = \overline{x}(y_d)$ for $y_d > y_r$. Next, $\phi(y_d)$ 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 A4. *Each $y_d \in X$ induces a unique policy lottery, with expectation $\mu_x(y_d)$ and variance $\sigma_x^2(y_d)$. If μ_x is weakly increasing on a compact interval $S \subseteq X$, then $U(y_d, y_p)$ satisfies the single-crossing condition on $S \times X$.*

Proof. Lemma 1 implies $\mu_x(y_d)$ and $\sigma_x^2(y_d)$ for all y_d. If $y_d' \geq y_d$, then Assumption 3 implies

$$U(y_d'; y_p) - U(y_d; y_p) = (\mu_x(y_d) - y_p)^2 - (\mu_x(y_d') - y_p)^2 + \sigma_x^2(y_d) - \sigma_x^2(y_d'),$$

so $\frac{\partial}{\partial y_p} U(y_d'; y_p) - \frac{\partial}{\partial y_p} U(y_d; y_p) \geq 0$ if and only if $\mu_x(y_d') - \mu_x(y_d) \geq 0$. Therefore if μ_x is increasing on a compact interval $S \subseteq X$, then U satisfies increasing differences on $S \times X$, which implies that it has the single-crossing property. □

Lemma A5. *There exist unique* $\underline{\pi} \in [\underline{x}_\ell, y_\ell)$ *and* $\overline{\pi} \in [y_r, \overline{x}_r)$ *such that* μ_x *is strictly increasing on* $[\underline{\pi}, \overline{\pi}]$*. Additionally, (i)* $\max\{\rho_L, \rho_R\} \leq \frac{1}{2\delta}$ *implies* $\underline{x}_\ell = \underline{\pi} < \overline{\pi} = \overline{x}_r$*, (ii)* $\rho_R > \frac{1}{2\delta}$ *implies* $\underline{x}_\ell < \underline{\pi} < \overline{\pi} = \overline{x}_r$*, and (iii)* $\rho_L > \frac{1}{2\delta}$ *implies* $\underline{x}_\ell = \underline{\pi} < \overline{\pi} < \overline{x}_r$*.*

Proof. For almost all y_d, we have $\mu_x'(y_d) = \rho_d + \rho_L \underline{x}'(y_d) + \rho_R \overline{x}'(y_d)$. First, Lemma 3 implies that $\underline{x}'(y_d) > 0$ and $\overline{x}'(y_d) > 0$ for all $y_d \in (y_\ell, y_r)$, so $\mu'(y_d) > 0$ on (y_ℓ, y_r). Next, $y_d \in (\underline{x}_\ell, y_\ell)$ implies $\underline{x}'(y_d) = -\overline{x}'(y_d)$ is strictly increasing and concave with $\lim_{y_d \to \underline{x}_\ell^+} \underline{x}'(y_d) = \frac{\delta \rho_d}{1 - \delta \rho_E}$ and $\lim_{y_d \to y_\ell^-} \underline{x}'(y_d) = 0$. Thus, $\rho_R \leq \frac{1}{2\delta}$ implies $\mu'(y_d) > 0$ on $(\underline{x}_\ell, y_\ell)$, so $\underline{\pi} = \underline{x}_\ell$; and $\rho_R > \frac{1}{2\delta}$ implies existence of a unique $\underline{\pi} \in (\underline{x}_\ell, y_\ell)$ such that $\mu'(y_d) < 0$ on $(\underline{x}_\ell, \underline{\pi})$ and $\mu'(y_d) > 0$ on $(\underline{\pi}, y_\ell)$. Then, a symmetric argument yields analogous characterization for $\overline{\pi}$. Finally, continuity of $\mu(y_d)$ implies μ_x is strictly increasing on $[\underline{\pi}, \overline{\pi}]$. □

Lemma A6. *The mapping* $\hat{y}_d(y_p) \equiv \operatorname*{argmax}_{y_d \in [y_\ell, y_r]} U(y_d; y_p)$ *is equivalent to an onto 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_p^**, (ii)* $y_p < y_p^*$ *implies* $\hat{y}_d(y_p) \in (y_p, y_p^*]$*, and (iii)* $y_p > y_p^*$ *implies* $\hat{y}_d(y_p) \in [y_p^*, y_p)$*.*

Proof. By Lemmas 3 and A3, for all $y_d \in (y_\ell, y_r)$ we have: $\underline{x}''(y_d) < 0 < \underline{x}'(y_d)$ and $0 < \min\{\overline{x}'(y_d), \overline{x}''(y_d)\}$.

Then, (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, Lemmas A4 and A5 imply that $U(y_d; y_p)$ satisfies the single-crossing condition on $[y_\ell, y_r] \times X$. Thus, $\hat{y}_d(y_p)$ is single-valued, continuous, increasing, and onto. 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 $\frac{\partial U(y_d;y_p)}{\partial y_d}\Big|_{y_d=y_p} = 0$. Define $\lambda : (y_\ell, y_r) \to \mathbb{R}$ as

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

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_p^* is unique. Furthermore, $y_p^* \in \{y_\ell, y_r\}$ if λ does not change sign, and otherwise $y_p^* \in (y_\ell, y_r)$. Finally, since $\hat{y}_d(y_p)$ weakly increasing and $\lambda'(y_p) < 0$, we know that (i) $y_p < y_p^*$ implies $\hat{y}_d(y_p) \in (y_p, y_p^*]$, and (ii) $y_p > y_p^*$ implies $\hat{y}_d(y_p) \in [y_p^*, y_p)$. □

Define $\tilde{y}_d : [0, y_\ell] \to [\underline{x}_\ell, y_\ell]$ as $\tilde{y}_d(y_p) \equiv \text{argmax}_{y_d \in [\underline{x}_\ell, y_\ell]} U(y_d, y_p)$.

Lemma A7. *Under Assumptions 3 and 4, $\tilde{y}_d$ is single-valued, continuous, increasing, and satisfies $\tilde{y}_d(y_p) \subset [\underline{\pi}, y_\ell]$ for all y_p. 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* $\text{argmax}_{y_d \in [y_r, \overline{x}_r]} U(y_d; y_p)$.

Proof. Fix $y_p \leq y_\ell$. To begin, we show that $\tilde{y}_d(y_p) \subset [\underline{\pi}, y_\ell]$. Lemma A5 implies μ is strictly decreasing on $[\underline{x}_\ell, \underline{\pi}]$ and strictly increasing on $[\underline{\pi}, y_\ell]$. The result is trivial if $\underline{\pi} = \underline{x}_\ell$, so suppose $\underline{\pi} > \underline{x}_\ell$ and consider $y_d, y_d' \in [\underline{x}_\ell, y_\ell]$ satisfying $y_d < y_d'$. Under Assumption 3,

$$U(y_d, y_p) - U(y_d', y_p) = (\mu_x(y_d) - y_p)^2 - (\mu_x(y_d') - y_p)^2 + \sigma_x^2(y_d) - \sigma_x^2(y_d').$$

By Lemma A5, we know that $\underline{\pi} > \underline{x}_\ell$ if and only if $\rho_R > \frac{1}{2\delta}$, which implies $\mu(y_d) > \mu(y_d') > y_r$.

Therefore $(\mu_x(y_d)-y_p)^2-(\mu_x(y_d')-y_p)^2>0$. Furthermore, $\sigma_x^2(y_d)>\sigma_x^2(y_d')$ since we have:

$$\begin{aligned}\sigma_x^2(y_d)-\sigma_x^2(y_d') = \rho_R[(\overline{x}(y_d)-\mu(y_d))^2-(\overline{x}(y_d')-\mu(y_d'))^2] \\ +\rho_L[(\underline{x}(y_d)-\mu(y_d))^2-(\underline{x}(y_d')-\mu(y_d'))^2] \\ +\rho_d[(y_d-\mu(y_d))^2-(y_d'-\mu(y_d'))^2] \\ +\rho_r[(y_r-\mu(y_d))^2-(y_r-\mu(y_d'))^2] \\ +\rho_\ell[(y_\ell-\mu(y_d))^2-(y_\ell-\mu(y_d'))^2],\end{aligned}$$

where (i) the first term is positive because $\overline{x}(y_d)>\mu(y_d)$ always holds and $\overline{x}'(y_d)-\mu'(y_d)=-\rho_d-\rho_L\underline{x}'(y_d)+(1-\rho_R)\overline{x}'(y_d)<0$ for all $y_d\in(\underline{x}_\ell,y_p)$, so $(\overline{x}(y_d)-\mu(y_d))^2>(\overline{x}(y_d')-\mu(y_d'))^2$; (ii) the second term is positive because $\underline{x}(y_d)<\underline{x}(y_d')<\mu(y_d')\le\mu(y_d)$; and (iii) the last three terms are positive because $y_d<y_d'\le y_\ell\le y_r<\mu(y_d')\le\mu(y_d)$. Altogether, this implies $U(y_d',y_p)>U(y_d,y_p)$. It follows that $U(\underline{\pi},y_p)>U(y_d,y_p)$ for all $y_d\in[\underline{x}_\ell,\underline{\pi})$. Thus, we have shown $\tilde{y}_d(y_p)=\text{argmax}_{y_d\in[\underline{\pi},y_\ell]}U(y_d,y_p)$.

Next, Lemmas A4–A5 and the theorem of the maximum imply that $\tilde{y}_d(y_p)$ must be non-empty, upper hemicontinuous, compact valued, and increasing. Furthermore, Lemma A2 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

$$\frac{\partial U(y_d;y_p)}{\partial y_d}\propto-\rho_d(y_d-y_p)-\underline{x}'(y_d)[\rho_L(\underline{x}(y_d)-y_p)-\rho_R(\overline{x}(y_d)-y_p)]=0 \qquad (14)$$

and

$$\frac{\partial^2 U(y_d;y_p)}{\partial y_d^2}\propto-\rho_d-\underline{x}''(y_d)[\rho_L(\underline{x}(y_d)-y_p)-\rho_R(\overline{x}(y_d)-y_p)]-[\underline{x}'(y_d)]^2\rho_E<0. \qquad (15)$$

Since $y_p<\tilde{y}_d(y_p)$ and $\underline{x}'(y_d)>0$, the first-order condition in (14) hold 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 unique 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 (14) 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 unique.

Additionally, $\tilde{y}_d(y_p)$ is strictly increasing on $[E_L, y_p]$ because (i) applying the implicit function theorem yields $\tilde{y}_d'(y_p) > 0$ if $\mu'(y_d) > 0$ at $y_d = \tilde{y}_d(y_p)$, and (ii) Lemma A5 shows that $\mu'(y_d) > 0$ on (π, y_ℓ).

Finally, $\tilde{y}_d(y_p) = \underline{x}_\ell$ for all $y_p < E_L$ because $\tilde{y}_d(y_p)$ is continuous and increasing.

Analogous arguments establish the result for $y_p \geq y_r$. $\square$

Proof of Proposition 2.

By Lemma A2, we know: (i) $y_d < E_L$ implies $y_d^*(y_d) = [0, \underline{x}_\ell]$; (ii) $y_p \in (E_L, y_\ell)$ implies $y_d^*(y_d) = \tilde{y}_d(y_p)$; and (iii) $y_p \in [y_\ell, y_r]$ implies $y_d^*(y_d) = \hat{y}_d(y_p)$. Thus, $y_d \in (\underline{x}(y_\ell), y_\ell)$, implies $y_d^*(y_d) \in \{\tilde{y}_d(y_p), \hat{y}_d(y_p)\}$. Lemmas A6 and A7 imply that $\hat{y}_d(y_p)$ is single-valued and $\tilde{y}_d(y_p)$ is single-valued on $(E_L, y_\ell]$. Additionally, Lemma A7 implies that: (i) $y_d^*(y_\ell) = y_\ell$ if and only if $\tilde{y}_d(y_\ell) = \hat{y}_d(y_\ell) = y_\ell$ and (ii) $\tilde{y}_d(y_p) < \hat{y}_d(y_p)$ for all $y_p < y_\ell$. Thus, there exists a $\underline{y}_p \in (\underline{x}(y_\ell), y_\ell]$ such that $y_d^*(\underline{y}_p) = \{\tilde{y}_p(\underline{y}_p), \hat{y}_p(\underline{y}_p)\}$. To show $\hat{y}_p$ is unique, it suffices to verify $y_d^*(y_p)$ is increasing. Lemma A7 implies

$$y_d^*(y_p) = \begin{cases} \operatorname{argmax}_{y_d \in [0,1]} \; U(y_p, y_d) & \text{if} \quad \max\{\rho_L, \rho_R\} \leq \frac{1}{2\delta} \\ \operatorname{argmax}_{y_d \in [\underline{\pi}, 1]} \; U(y_p, y_d) & \text{if} \quad \rho_R > \frac{1}{2\delta} \\ \operatorname{argmax}_{y_d \in [0, \overline{\pi}]} \; U(y_p, y_d) & \text{if} \quad \rho_L > \frac{1}{2\delta}. \end{cases}$$

Then, Lemmas A4 and A5 imply $y_d^*(y_p)$ is increasing, so $\underline{y}_p$ is unique. Lemmas A6 and A7 then imply y_d^* is single-valued on $(E_L, y_r] \setminus \underline{y}_p$ and strictly increasing on $(E_L, \underline{y}_p)$. An analogous argument establishes the results for $y_d \geq y_\ell$. $\square$

B.2 Proof of Proposition 3.

First, Propositions 1 and 2 imply that any fixed point of y_d^* must be in $[y_\ell, y_r]$, where $y_d^*(y_p) = \hat{y}_d(y_p)$. Second, by Lemma A6: $y_p^*|_{[y_\ell, y_r]}$ has a unique fixed point; $y_p < y_p^*$ implies $\hat{y}_d(y_p) \in (y_p, y_p^*]$; and $y_p > y_p^*$ implies $\hat{y}_d(y_p) \in [y_p^*, y_p)$. Third, by Proposition 2: $y_p \in (E_L, \underline{y}_p)$

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_p^* is the unique fixed point of $y_d^*|_{(E_L,E_R)}$. $\square$

B.3 Proof of Proposition 4.

By Lemma A2, (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 $y_p^* = y_\ell$ and (ii) $\overline{y}_p = y_r$ if and only if $y_p^* = y_\ell$. Therefore uniqueness of y_p^* implies $y_r \in \Delta$ or $y_\ell \in \Delta$. The characterization using λ follows directly from the characterization of y_p^* in Lemma A6. $\square$

B.4 Proof of Proposition 5.

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 $y_p^* = y_\ell$, (ii) $\lambda(y_r) \geq 0$ implies $y_p^* = y_r$, and (iii) otherwise $\lambda(y_p^*) = 0$ at $y_p^* \in (y_\ell, y_r)$. Thus, y_p^* is weakly increasing.

2. From Lemma A7, E_L is the smallest $y_p < \underline{x}_\ell$ such that $\left.\frac{\partial U(y_d,E_L)}{\partial y_d}\right|_{y_d=\underline{x}_\ell^+} \geq 0$. Then, computation yields $\left.\frac{\partial^2 U(y_d,E_L)}{\partial y_d \partial\rho_L}\right|_{y_d=\underline{x}_\ell^+} - \left.\frac{\partial^2 U(y_d,E_L)}{\partial y_d \partial\rho_R}\right|_{y_d=\underline{x}_\ell^+} = -\frac{\delta\rho_d[(\underline{x}_\ell - E_L)+(\overline{x}_r - E_L)]}{1-\delta\rho_E} < 0$, so E_L weakly increases in ρ_L. By an analogous argument, E_R weakly increases in ρ_L.

3. Recall that $\underline{y}_p = y_\ell$ if and only if $y_\ell = y^*$. Since $\frac{\partial\lambda(y_d)}{\partial\rho_L} - \frac{\partial\lambda(y_d)}{\partial\rho_R} > 0$ for all $y_d \in (y_\ell, y_r)$, there is a unique ρ_L' such that $y^* = y_\ell$ if and only if $\rho_L \leq \rho_L'$. Thus $\underline{y}_p = y_\ell$ for all $\rho_L \leq \rho_L'$ and $\underline{y}_p < y_\ell$ otherwise. To complete the proof, we show that $\underline{y}_p$ is decreasing

on $\rho_L \in (\rho'_L, 1)$. Then, $\underline{y}_p$ is the unique $y_p < y_\ell$ satisfying

$$
\begin{aligned}
0 &= U(\hat{y}, y_p) - U(\tilde{y}, y_p) && (16)\\
&= \rho_L \Big(-(\underline{x}(\hat{y}) - y_p)^2 + (\underline{x}(\tilde{y}) - y_p)^2\Big) + \rho_R \Big(-(\overline{x}(\hat{y}) - y_p)^2 + (\overline{x}(\tilde{y} - y_p)^2\Big) \\
&\quad + \rho_d \Big(-(\hat{y} - y_p)^2 + (\tilde{y} - y_p)^2\Big), && (17)
\end{aligned}
$$

where (17) follows by definition. Moreover, $U(\hat{y}, y_p) - U(\tilde{y}, y_p) > 0$ for all $y_p \in (\underline{y}_p, y_\ell)$ and $U(\hat{y}, y_p) - U(\tilde{y}, y_p) < 0$ for all $y_p \in (\underline{x}_\ell, \underline{y}_p)$. Therefore $U(\hat{y}, y_p) - U(\tilde{y}, y_p)$ is increasing at $y_p = \underline{y}_p$. So $\underline{y}_p$ must be decreasing in ρ_L if $U(\hat{y}, y_p) - U(\tilde{y}, y_p)$ is increasing in ρ_L at $y_p = \underline{y}_p$.

Letting $\xi(y) = \frac{\partial U(y, y_p)}{\partial \rho_L} - \frac{\partial U(y, y_p)}{\partial \rho_R}$, the envelope theorem implies

$$
\xi(\hat{y}) - \xi(\tilde{y}) = (\underline{x}(\tilde{y}) - y_p)^2 - (\underline{x}(\hat{y}) - y_p)^2 + (\overline{x}(\hat{y}) - y_p)^2 - (\overline{x}(\tilde{y}) - y_p)^2. \qquad (18)
$$

To show a contradiction, suppose $\xi(\hat{y}) - \xi(\tilde{y}) < 0$. Then, we must have

$$
\begin{aligned}
0 &= U(\hat{y}, y_p) - U(\tilde{y}, y_p) \\
&< \rho_d[-(\hat{y} - y_p)^2 + (\tilde{y} - y_p)^2] + \rho_E[-(\underline{x}(\hat{y}) - y_p)^2 + (\underline{x}(\tilde{y}) - y_p)^2], && (19)
\end{aligned}
$$

where (19) follows from (17) because (i) $U(\hat{y}, y_p) - U(\tilde{y}, y_p)$ is increasing in $(\overline{x}(\tilde{y}) - y_p)^2$, and (ii) $0 > \xi(\hat{y}) - \xi(\tilde{y})$ implies $(\overline{x}(\tilde{y}) - y_p)^2 > (\underline{x}(\tilde{y}) - y_p)^2 - (\underline{x}(\hat{y}) - y_p)^2 + (\overline{x}(\hat{y}) - y_p)^2$. Thus, since adding and subtracting $\rho_R[-(\underline{x}(\hat{y}) - y_p)^2 + (\underline{x}(\tilde{y}) - y_p)^2]$ in (17) yields

$$
\begin{aligned}
U(\hat{y}, y_p) - U(\tilde{y}, y_p) = {}& \rho_R[-(\overline{x}(\hat{y}) - y_p)^2 + (\overline{x}(\tilde{y}) - y_p)^2 + (\underline{x}(\hat{y}) - y_p)^2 - (\underline{x}(\tilde{y}) - y_p)^2] \\
&+ \rho_d[-(\hat{y} - y_p)^2 + (\tilde{y} - y_p)^2] + \rho_E[-(\underline{x}(\hat{y}) - y_p)^2 + (\underline{x}(\tilde{y}) - y_p)^2],
\end{aligned}
$$

we know (19) holds if and only if

$$
\begin{aligned}
0 &> \rho_R[-(\overline{x}(\hat{y}) - y_p)^2 + (\overline{x}(\tilde{y}) - y_p)^2 + (\underline{x}(\hat{y}) - y_p)^2 - (\underline{x}(\tilde{y}) - y_p)^2] \\
&\propto \xi(\tilde{y}) - \xi(\hat{y}).
\end{aligned}
$$

Therefore we must have $\xi(\hat{y}) - \xi(\tilde{y}) < 0 < \xi(\hat{y}) - \xi(\tilde{y})$, a contradiction. $\square$

C Extensions

C.1 Proof of Proposition 6.

By Lemma A5, $\max\{\rho_L, \rho_R\} \leq \frac{1}{2\delta}$ implies μ_x is increasing on X. Thus, Lemma A4 implies $U(y_d, y_p)$ satisfies the single-crossing property on X^2. $\square$

C.2 Proof of Proposition 7.

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

C.3 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 (9). The result for $y_b(y_a)$ is analogous. $\square$

C.4 Proof of Proposition 8.

Corollary 1.2 and Proposition 2 together imply $y_a^* \in (y_{p_a}, y_M)$ and $y_b^* \in (y_M, y_{p_b})$. Lemma 5 implies uniqueness. $\square$

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* (4), 1181–1206.

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.

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.

Kartik, Navin, SangMok Lee, and Daniel Rappoport, "Single-Crossing Differences in Convex Environments," *Working Paper*, 2022.

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.

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.

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.

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

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.