Dynamic Policy Competition, Ideological Polarization and the Value of Veto Rights *

Salvatore Nunnari
Columbia University
snunnari@columbia.edu

Jan Zápal
IAE-CSIC, Barcelona GSE and CERGE-EI Prague
jan.zapal@iae.csic.es

January 8, 2014

Abstract

We study an environment where two parties alternate in office and the policies they propose have to be approved by a veto player. We propose two measures to describe the degree of conflict among agents. We define parties’ antagonism as the ideological disagreement between parties, and parties’ extremism as the ideological disagreement between each party and the veto player. These two measures do not coincide when parties care about multiple issues. We show that forward-looking incumbents have an incentive to implement policies favored by the veto player, in an attempt to constraint future incumbents. This incentive grows as antagonism increases, with policies moving closer to the ideal point of the veto player. On the other hand, as extremism increases, the short term incentives dominate and policies move away from the ideal point of the veto player. We discuss the empirical implications for the debate on the polarization of American politics.

JEL Classification: C72, C73, D72, D78
Keywords: Political Polarization; Dynamic Policy Competition; Multiple Issues; Veto Power.

* We are grateful to Enrique Aragones, Aislinn Bohren, Torun Dewan, Daniel Diermeier, John Duggan, Jon Eguia, Justin Grimmer, Esther Hauk, Ana McDowall, Massimo Morelli, Pablo Montagnes, Hannes Mueller, Clara Ponsati, Carlo Prato, Milan Svolik and seminar participants at IAE-CSIC, the Conference in Honor of Norman Schofield at Washington University in St Louis, the 2013 EPSA Conference in Barcelona, the CRENoS IV Workshop on Institutions, Individual Behavior and Economic Outcomes in Alghero, and the 2013 APSA Annual Meeting in Chicago for helpful comments and interesting discussions. Jan Zapal acknowledges financial support from the Post-Doc Research Fund of the Charles University in Prague. All remaining errors are our own.
1 Introduction

A large body of recent research in political science has been devoted to identifying and explaining ideological polarization, especially but not only in the United States. There is strong evidence that the U.S. Congress has grown progressively more polarized since the 1970s (McCarty, Poole and Rosenthal 2006), as well as some evidence of a polarization trend in presidential platforms (Budge, Klingemann, Volkens, Bara and Tanenbaum 2001, Klingemann, Volkens, Bara, Budge and McDonald 2006). Many empirical studies link this increasing ideological divide between the main American parties to legislative gridlock, elite incivility, income inequality, and voter disengagement (Layman and Carsey 2002b, Fiorina and Abrams 2008, Hetherington 2001). Across a broader range of countries, polarization is associated with democratic breakdown, corruption, and economic decline (Brown, Touchton and Whitford 2011, Frye 2002, Linz and Stepan 1978).

To complement these empirical findings, many existing models of electoral competition and policymaking show that parties with conflictual preferences are socially suboptimal (Persson and Svensson 1989, Alesina and Tabellini 1990, Azzimonti 2011). However, as we hope to show, these conclusions depend crucially on a number of stylized assumptions about the nature of the disagreement among political actors. While we lack a precise notion of ideological polarization when parties care about multiple dimensions, how we measure this concept crucially affects the conclusions on its welfare consequences.

In this article, we develop a model designed to study the policy consequences of ideological conflict and the different value of agenda setting and veto rights, in a dynamic bargaining setting where the location of the current status quo policy is determined by the policy implemented in the previous period. This is an important feature of many policy domains—for instance personal income tax rates or entitlement programs—where legislation remains in effect until the legislature passes a new law. In each of an infinite number of periods, one of two parties with conflicting preferences controls the legislature and the majority party leader has the power to propose the adoption of a bi-dimensional policy. The proposed policy is implemented if it receives the approval of an institutional veto player (an executive body, the judiciary system, or the median legislator on the legislative assembly floor). Otherwise, the status quo policy prevails and the policy implemented is the same as in the previous period. In this sense, the status quo policy evolves endogenously.

We highlight that—in a multidimensional environment—we can use many different
measures to describe the degree of conflict—of polarization—among agents’ preferences in the bargaining environment and we propose two such measures. The first measure, which we label extremism, is the ideological distance of each party from the veto player (or a moderate, decisive voter in the legislature). The second measure, which we call antagonism, is the ideological distance that separates the two parties from each other and summarizes the degree of political competition between the agenda setters. These two measures coincide in a one dimensional policy space, where the ideological distance between the two parties can increase only as they move further away from the veto player. However, they do not coincide in a two-dimensional setting: here, the two parties can be very close (when they share views on both dimensions) or very different (when they are perfectly opposed in one dimension), without altering their overall distance from the veto player.

There are three main questions that we wish to address with this simple setting. First, what is the impact of how proposal and veto rights are allocated across actors in a dynamic bargaining situation, where there is competition over agenda setting and policies evolve endogenously over time? Second, what are the conditions on ideological disagreement under which one of these prerogatives is more valuable to a political actor than another? And third, how does the allocation of proposal and veto rights, together with ideological disagreement, affect the electorate’s welfare?

We fully characterize a Stationary Markov Perfect Equilibrium (SMPE) of the dynamic policy competition game described above and prove it exists for any discount factor, any initial status quo, any stochastic process of power switch, and any degree of antagonism and extremism. According to the results of our model, long run policies tends to be more moderate: (i) the larger is the degree of antagonism between alternating governments; (ii) the smaller is parties’ extremism; (iii) the larger is parties’ patience; (iv) the larger is the difference in incumbency advantage between the two parties. The key idea is that an incumbent government that is uncertain about whether it will be re-elected has incentives to behave strategically by implementing policies that restrict the choice set of future governments. The more the preferences of the incumbent government depart from the preferences of potential future governments, the more the incumbent will try to restrict future governments’ choice set. The degree of conflict between the parties and the veto holders, on the other hand, does not affect the strategic incentives

---

1 The only general existence result for dynamic bargaining games applies to settings with stochastic shocks to preferences and the status quo (Duggan and Kalandrakis 2012). As these features are not present in our model, proving existence is a necessary step of the analysis.
to moderate.

More generally, our paper suggests that the influence veto players exert on long run policy outcomes is a function of the degree of antagonism and extremism of the political system. In particular, increasing the degree of extremism reduces the welfare of the veto holder, while increasing antagonism increases the influence exerted by the veto holder on long run policy outcomes. Our analysis provides novel and testable empirical implications: the differences in the degree and dimensionality of political conflict can contribute to explain the variance in the impact different political institutions and actors have over outcomes in different countries, or in the same country at different points in time, as well as the volatility of the observed policies. Moreover, when we consider a moderate median voter who shares policy preferences with the institutional veto player, or reinterpret our bargaining model as a model of direct democracy, our model shows that antagonism can have counter-intuitive welfare implications: an electorate that prefers moderate policies is best served by highly antagonist political elites that are perfectly opposed on one dimension.

This paper contributes primarily to the theoretical literature on the consequences of veto power in legislatures. A large number of studies build on models of legislative bargaining à la Baron and Ferejohn (1989) or agenda setting à la Romer and Rosenthal (1978) to examine the role of veto power in policy making. Most of these papers model specific environments and focus on the case of the U.S. Presidential veto (Matthews 1989, Diermeier and Myerson 1999, Cameron 2000, McCarty 2000, Groseclose and McCarty 2001, Primo 2006, Duggan, Kalandrakis and Manjunath 2008). A common limitation of this literature, and the main point of departure with our paper, is the focus on static settings: the legislative interaction ceases once the legislature has reached a decision, and policy cannot be modified after its initial introduction. In these frameworks, any conclusion on the effect of veto power on policy outcomes depends heavily on the specific assumptions on the status quo policy and the proposer has a significant bargaining advantage (Krehbiel 1998, Tsebelis 2002). In our paper, the status quo policy is not exogenously specified but is rather the product of policy makers’ past decisions. More closely related to this paper, Nunnari (2012) shows that, in a dynamic divide-the-dollar game with an endogenously evolving status quo policy, the veto holder can steer the policy towards his ideal outcome. Our paper differs from this work in two fundamental

---

[^2]: In this sense, the present study belongs to a growing literature on legislative policy making with an endogenous status quo and farsighted players (Anesi and Seidmann, 2012; Battaglini and Palfrey, 2012; Baron, 1996; Bowen, Chen and Eraslan, 2012; Diermeier and Fong, 2011; Dziuda and Loeper, 2012; Forand, 2010; Kalandrakis, 2004).
respects: first, in our paper, the veto player has no power to propose—a faculty that plays a crucial role in the equilibrium constructed by Nunnari (2012); and second, the bargaining is not over how to divide a fixed amount of resources with perfectly conflicting preferences, but over the location of a policy in a two-dimensional space, with concave preferences that can be more or less opposed. These modeling differences allow us to investigate the relationship between preferences’ divergence and veto right leverage, and to pose the question of the welfare consequences of veto power and political polarization.

Finally, our work is related to political economic models that directly address the relationship between elite polarization and policy outcomes. In most models, conflictual political preferences are socially suboptimal (Persson and Svensson 1989, Alesina and Tabellini 1990, Azzimonti 2011, Prato 2013). Our paper shares with this literature the idea that forward-looking incumbents have incentives to strategically position current policies to affect future political outcomes, and that these incentives are stronger when the conflict of preferences is starker. On the other hand, in our model disagreement can be over multiple dimensions (rather than over the desired level of public spending or the public good to invest in) and the channel to constraint a future incumbent is the demands of the veto player (or the electorate) rather than inefficient or misdirected spending. Our novel approach shows that—when the ideological conflict is multidimensional—preference divergence does not necessarily lead to higher inefficiencies and welfare losses for the electorate, but it could have the opposite effect.

More recent studies argued that polarized parties and divergent platforms can be welfare enhancing (Bernhardt, Duggan and Squintani 2009, Van Weelden 2013). In these works, polarization coincides with our notion of extremism and helps the electorate through channels that are different than the one highlighted in our paper. Van Weelden (2013) shows that, in a more polarized political environment, the incumbent gives up rent extraction for fear of being replaced by a challenger with markedly different policy preferences. Bernhardt, Duggan and Squintani (2009) analyze a static and unidimensional electoral competition where more polarized parties propose more extreme electoral platforms; for moderate levels of polarization, this is welfare enhancing because it provides the median voter with a more varied choice. In our case, on the other hand, the beneficial effect of antagonism comes from more moderate policies (implemented strategically by forward-looking incumbents), and increased extremism is always detrimental to the veto player.

The article proceeds as follows. Section 2 gives a detailed presentation of the leg-
islative setup and introduces the equilibrium notion. Section 3 outlines the equilibrium analysis and gives the main results. Section 4 introduces a notion of ex-ante welfare of the veto player, whose comparative static is presented in Section 5. Section 6 proposes some extensions and alternative interpretations of the model. Section 7 concludes with a discussion of the empirical implications of our analysis.

2 Model and Equilibrium Notion

We describe here a dynamic model of policy making by ideological parties that, when in office, can set the agenda and propose a change of the status quo policy. In each period $t$ of an infinite horizon, one of two parties, 1 and 2, is selected to be the incumbent office holder, $I_t \in \{1, 2\}$. The identity of the incumbent in period $t$ is determined by a Markov process governed by the following recognition probabilities:

$$
P[I_t = i|I_{t-1} = i] = r_i
$$

$$
P[I_t = -i|I_{t-1} = i] = 1 - r_i
$$

where $r_i \in [0, 1]$, and, for party $i \in \{1, 2\}$, $-i = \{1, 2\} \setminus \{i\}$ denotes the opposing party. This process is fairly general and it encompasses many power shifting processes widely used by the literature as special cases. For example, with $r_1 = 1 - r_2$, the identity of the incumbent in period $t$ is independent of the identity of the incumbent in $t - 1$; with $r_1 = r_2 = 0$, the two parties alternate in power; and for $(r_1, r_2) \gg (0, 0)$ both parties enjoy an incumbency advantage, with inertia in the identity of the agenda setter.

The Policy Making Process. The incumbent party has the power to propose a bi-dimensional policy $p = (p^1, p^2) \in X \subseteq \mathbb{R}^2$. The proposed policy $p$ is implemented if and only if an institutional veto player, $v$, accepts it. If $v$ rejects the proposal, the policy implemented in the period is the status quo policy, $q \in X$, inherited from the previous period. The policy implemented in a period becomes the status quo policy in the next period and, as such, represents a dynamic linkage between periods.

---

3 The policy space $X$ can be either $\mathbb{R}^2$ or a compact, convex, and proper subset of $\mathbb{R}^2$ that is large enough to include the ideal policies of all the agents, as described below. Restricting attention to such a subset, $X \subset \mathbb{R}^2$, is without loss of generality.
Stage Utilities. The stage utility player \( i \in \{1, 2, v\} \) receives from policy \( p \in X \) is measured by the squared distance of \( p \) from \( i \)’s bliss point, or ideal policy, \( b_i = (b_1^i, b_2^i) \):

\[
u_i(p) = -(p^1 - b_1^i)^2 - (p^2 - b_2^i)^2
\]

Denoting by \( d(x, y) \) the usual Euclidean distance between \( x \in X \) and \( y \in X \), we can rewrite equation (1) as \( u_i(p) = -d^2(p, b_i) \). We will abuse notation slightly and will denote by \( d(x) = ||x|| \) the Euclidean norm (that is, the distance from the origin) of \( x \in X \). The utility \( i \) derives from a sequence of policies \( P = \{p_0, p_1, \ldots\} \) is the discounted sum of payoffs from each period:

\[
U_i(P) = \sum_{t=0}^{\infty} \delta_t^i u_i(p_t)
\]

where \( \delta_i \in [0, 1) \) is player \( i \)’s discount factor. We assume \( \delta_1 = \delta_2 = \delta \).

Parties’ Antagonism and Extremism. The model is shift and rotation invariant, so, without loss of generality, we define the veto player’s bliss point as the origin of the plane, \( b_v = (0, 0) \), party 1’s bliss point as \( b_1 = (b, 0) \) and party 2’s bliss point as \( b_2 = (b_1^2, b_2^2) \). We assume that parties’ bliss points are at the same distance from the veto player’s bliss point, that is, \( d(b_2) = b \). The parameter \( b \), thus, captures the ideological distance between the veto player and the two parties. We call this the degree of extremism of the parties.

A second, separate measure of ideological divergence is given by the angle \( \alpha \in [0, \pi] \) formed by the two vectors \( b_1 \) and \( b_2 \). This parameter captures how different the parties’ bliss points are from each other, regardless of their distance from the veto player’s ideal policy. When \( \alpha = 0 \), the parties’ bliss points are on the same ray departing from the origin (and coincide when parties are symmetric, or \( d(b_2) = b \), as we assume throughout the paper). When \( \alpha \in (0, \pi) \) the two parties diverge on both dimensions, keeping the distance from the origin of the plane constant and symmetric. When \( \alpha = \pi \), the two parties are perfectly opposed on one dimension and share the same ideology on the

\[\text{4 We use quadratic Euclidean preferences primarily for convenience. In Appendix A2 we present a model with general utility functions that are continuous, decreasing and weakly concave in } d(x, y), \text{ and prove results analogous to the ones presented below.}
\[\text{5 This is not without loss of generality. We discuss the case of parties with asymmetric bliss points in the Extensions and Discussion section below.}
\[\text{6 Notice that } \cos \alpha = \frac{b_1 \cdot b_2}{||b_1|| ||b_2||} \text{ where } b_1 \cdot b_2 \text{ is the usual inner product.} \]
second dimension. We call $\alpha$ the degree of antagonism of the parties.

The equilibrium strategies we characterize below will depend solely on $\alpha$ and on the distance between the status quo policy and the bliss point of the veto player (the origin of the plane). We denote with $k(x) = \frac{d(x)}{b} \geq 0$ the distance of policy $x \in X$ from the origin, relative to $b$. With this notation, $k(q)b_i$ is a point on the line connecting $b_v$ with $b_i$, at the the same distance from $b_v$ as the status quo policy $q$. Figure 1 shows the basic model parameters: a set of arbitrary bliss points for the three players ($b_v, b_1,$ and $b_2$) and the indifference curves generated by their Euclidean preferences over policies; the corresponding degree of antagonism ($\alpha$) and extremism ($b$); the status quo policy ($q$) and its distance from the origin ($d(q)$); and a point on the line connecting $b_2$ with the origin, at the same distance from the origin as the status quo policy ($k(q)b_2$). Figure 2 shows examples of parties’ bliss points for two different degrees of antagonism ($\alpha'>\alpha$) and two different degrees of extremism ($b'>b$).

Figure 1: Basic Model Parameters

**Strategies.** We focus on equilibria in pure Markov strategies (Maskin and Tirole, 2001). We assume that both the decision of the incumbent regarding which policy to propose, and the decision of the voter to accept the proposed policy depend solely on the status quo policy $q$. Markovian strategies that abstract from the history of play are standard in dynamic models of political economy (Baron 1996, Battaglini and Palfrey 2012, Battaglini, Nunari and Palfrey 2012, Duggan and Kalandrakis 2012, Kalandrakis 2004, Kalandrakis 2010, Forand 2010), capture the simplest form of behavior consistent with rationality, and clearly isolate the underlying strategic motives shaping the policy
competition between the two parties in a dynamic environment, independent of the time horizon. Additionally, the two parties interact over a long time horizon and can be represented by different politicians in different points in time. Therefore, strategies that potentially depend on events from the (distant) past and require coordination might be excessively demanding and inappropriate for the context at hand.

**Definition 1.** A Stationary Markov strategy for incumbent party \( i \in \{1, 2\} \) is a function \( \sigma_i : X \rightarrow X \), mapping each status quo policy \( q \) into a policy proposal \( p \). A Stationary Markov strategy for the veto player \( v \) is a function \( \sigma_v : X \times X \rightarrow \{\text{Yes}, \text{No}\} \) that maps each pair composed of a status quo policy \( q \) and incumbent’s proposal \( \sigma_i(q) = p \) into the decision to accept or reject.

**Expected Utilities.** Denote by \( V^i_j(q|\sigma) \) the expected utility of agent \( i \in \{1, 2, v\} \) from the infinite sequence of policies generated by the profile of strategies \( \sigma = (\sigma_1, \sigma_2, \sigma_v) \), at the beginning of a period and before the identity of the party in power is determined, when the incumbent in the previous period was \( j \in \{1, 2\} \), and the status quo is \( q \). Formally, if \( P(j, \sigma, q) = \{p_0, p_1, \ldots\} \) is a path of policies generated by play according to \( \sigma \), starting from the status quo \( q \) with incumbent \( j \), we have:

\[
V^i_j(q|\sigma) = \mathbb{E}[U_i(P(j, \sigma, q))] = \mathbb{E} \left[ \sum_{t=0}^{\infty} \delta^t u_i(p_t) \right]
\]  

**Equilibrium Notion.** In equilibrium, when party \( i \in \{1, 2\} \) is in power and chooses which policy to propose, it will consider only those policies that \( v \) accepts and will propose the acceptable policy \( p^* \) that maximizes its expected utility \( u_i(p) + \delta V^i_j(p|\sigma) \).
For ease of exposition and without loss of generality, we assume that, when \( p^* = q \), then \( \sigma_i(q) = q \).
Additionally, to avoid any open set complications, we assume that \( v \) approves \( p \) when she is indifferent between the proposal and the status quo. These two assumptions imply that any \( p \) proposed by the party in power is accepted and hence we do not need to track separately policies that are proposed and accepted.

**Definition 2.** A Stationary Markov Perfect Equilibrium (SMPE) is a profile of Stationary Markov strategies \( \sigma^* = (\sigma_1^*, \sigma_2^*, \sigma_v^*) \) such that, for any \( q \in X \),

\[
\sigma_i^* \in \arg \max_{\sigma_i} u_i(\sigma_i(q)) + \delta V_i^i(\sigma_i(q)|\sigma_i, \sigma_{-i}^*, \sigma_v^*)
\]

for \( i \in \{1, 2\} \), and given the incumbent’s proposal, \( \sigma_v^*(p, q) = yes \) if and only if

\[
u_v(p) + \delta_v V_v^i(p|\sigma^*) \geq u_v(q) + \delta_v V_v^i(q|\sigma^*)
\]

### 3 Equilibrium Analysis

#### 3.1 Simple Strategies

In the remainder, we will focus on a class of SMPE of the dynamic policy competition game, where the two parties use proposal strategies of a simple form, captured by a single parameter \( \hat{k}_i \).

**Definition 3.** A Simple (Stationary Markov Perfect) strategy \( \sigma_i \) for \( i \in \{1, 2\} \) satisfies

\[
\sigma_i(q) \equiv p_i(q) = \begin{cases} 
  k(q)b_i & \text{for } k(q) \leq \hat{k}_i \\
  \hat{k}_i b_i & \text{for } k(q) \geq \hat{k}_i
\end{cases}
\]

A Simple Stationary Markov Perfect Equilibrium (SSMPE) is a SMPE where the parties use simple strategies.

As discussed above, \( k(q) \) measures the distance of the status quo \( q \) from the origin of the plane. According to these simple strategies, when the status quo \( q \) is closer than \( \hat{k}_i \) to the origin, the incumbent party \( i \in \{1, 2\} \) proposes \( k(q)b_i \). This policy is located on the ray that connects \( b_v \) with \( b_i \), and it is at the same distance from \( b_p \) than the status.

---

7 When \( q \) maximizes \( i \)’s expected utility, \( i \) can obtain utility from \( q \) either directly by proposing \( q \) or indirectly by proposing a policy that will be rejected. We assume \( i \) does the former.
quo q. When instead, the status quo q is further than \( \hat{k}_i \) from the origin, the incumbent party \( i \in \{1, 2\} \) proposes \( k_i b_i \). This policy is on the same ray that connects \( b_v \) with \( b_i \), but it is \( \hat{k}_i \) distant from the origin. Figure 3 shows the policies corresponding to these simple strategies: for a number of arbitrary status quo policies, the arrow indicates the policy implemented by party 2 when he has the power to set the agenda. It is interesting to consider how the equilibrium policy evolves as \( k(q) \) increases (that is, as the status quo policy is further away from the origin): for low values of \( k(q) \), the equilibrium policy \( p_i(q) \) increases linearly (in the sense of getting further away from the origin) as \( k(q) \) increases; once \( k(q) \) reaches \( \hat{k}_i \), \( p_i(q) \) stays constant and it is not affected by a further increase of \( k(q) \).

Figure 3: Simple Stationary Markov Perfect Proposal strategy

![Figure 3](image1.png)

3.2 Results

Proposition 1 below shows that a SSMPE of the dynamic policy competition game exists and is generally unique, and it fully characterizes it.

**Proposition 1** (SSMPE of Dynamic Policy Competition Game). Assume, without loss of generality, that \( r_1 \geq r_2 \). Then:

1. \( \hat{k}_1 = 1 \) and \( \hat{k}_2 = \max \left[ 0, \frac{1 - \delta r_1 + \delta(1 - r_2) \cos \alpha}{1 - \delta r_1 + \delta(1 - r_2)} \right] \) characterize a SSMPE;

2. if \( r_1 > r_2 \), this SSMPE is unique; if \( r_1 = r_2 \), there exists exactly one additional ‘mirror’ SSMPE where \( k_1 \) are \( \hat{k}_2 \) reversed;
3. in any SSMPE, $\sigma_v(p, q) = \text{yes if and only if } d(p) \leq d(q)$;

4. starting from $q$, if $r_1 < 1$, or if $r_1 = 1$ and party 2 is in power in the initial period, SSMPE policies converge to alternating between

   (a) $k_2b_1$ and $k_2b_2$, if $k(q) \geq \hat{k}_2$
   
   (b) $k(q)b_1$ and $k(q)b_2$, if $k(q) \leq \hat{k}_2$

   depending on the identity of the incumbent;

5. starting from $q$, if $r_1 = 1$ and party 1 is in power in the initial period, SSMPE policies converge to $b_1$ if $k(q) \geq 1$ or $k(q)b_1$ if $k(q) < 1$.

Proof. See Appendix A1.

We now explain the basic intuition behind this equilibrium. Recall that the expected utility of agent $i \in \{1, 2, v\}$ from policy $p \in X$, when party $j \in \{1, 2\}$ is in power, is $u_i(p) + \delta_i V_i^j(p|\sigma)$. In this expression, the first term captures the current utility, and the second term captures the future utility.

We start with the veto player. We know from Definition 2 that, in any SMPE (not necessarily in simple strategies), the veto player prefers policy $p$ to the status quo $q$ if and only if $u_v(p) + \delta_v V_v^1(p|\sigma) \geq u_v(q) + \delta_v V_v^2(q|\sigma)$. Consider first the case of a perfect myopic veto player who has no concern about the future and only cares about her utility at the end of the current period. In this case, the veto player compares $u_v(p)$ and $u_v(q)$, and only accepts those policies that are at least as close to her bliss point as the status quo or, formally, those policies $p$ such that $d(p) \leq d(q)$. These policies lie in the circle with center at $b_v = (0, 0)$ and radius $d(q)$.

Now consider a veto player who is forward looking and cares both about her current utility and how the present policy affects her future stream of payoffs. In a SMPE in simple strategies, a forward-looking veto player has the same strategic incentives as a myopic voter. This is because the veto player’s value function, $V_v^1(p|\sigma)$, is weakly decreasing in the distance of the policy from the origin, $d(p)$, just as the stage utility, $u_v(p)$. As a consequence, the veto player’s discount factor $\delta_v$ does not play any role in the SSMPE from Proposition 1: the proposing and voting behavior of the agents does not depend on the degree of patience of the veto player.

\footnote{Figure 4 outlines two such acceptance sets (circles) for two illustrative status quo policies $q_1$ and $q_2$. For given status quo policy, the veto player accepts any policy on the boundary or strictly inside the appropriate circle.}
We now explain the equilibrium behavior of the parties with the help of Figure 4. We first note that in a SSMPE the policies proposed by party $i \in \{1, 2\}$ lie always on the ray starting at $b_v$, the origin of the plane, and passing through $b_i$. We call such a ray a $b_i$-ray. We can understand why this is the case by investigating the expected utility $i$ derives from proposing $\mathbf{p}$, $u_i(\mathbf{p}) + \delta V_i(\mathbf{p}|\sigma)$. The second term, the discounted value function capturing the future stream of payoffs, depends on $d(\mathbf{p})$ but not on the specific location of $\mathbf{p}$. This is because the strategies of all the players depend on the distance of the status quo from the origin, but not on its exact location. As a result, moving $\mathbf{p}$ along any circle centered at $b_v$, the expected utility of $i$ increases as $\mathbf{p}$ approaches the $b_i$-ray, because this increases the utility accrued at the end of the period but maintains constant the expected future utility.

Figure 4 shows many different circles centered at $b_v$, in dashed lines. The exact location of $\mathbf{q}_1$ and $\mathbf{q}_2$ in the Figure is arbitrary. However, the distance of these two points from $b_v$ is fully determined by Proposition 1: $k(\mathbf{q}_1) = \hat{k}_1 = 1$ and $k(\mathbf{q}_2) = \hat{k}_2$. The simple strategies from Proposition 1 prescribe that party 2 proposes $\hat{k}_2 \mathbf{b}_2$ for any status quo $\mathbf{q}$ with $d(\mathbf{q}) \geq d(\mathbf{q}_2)$ and $k(\mathbf{q}) \mathbf{b}_2$ for any status quo $\mathbf{q}$ with $d(\mathbf{q}) \leq d(\mathbf{q}_2)$. Similarly, party 1 proposes $\hat{k}_1 \mathbf{b}_1 = \mathbf{b}_1$ or $k(\mathbf{q}) \mathbf{b}_1$ depending on whether $d(\mathbf{q}) \geq d(\mathbf{q}_1)$.

We first discuss the intuition behind party 2’s strategy. When $d(\mathbf{q}) \leq d(\mathbf{q}_2)$, proposing $k(\mathbf{q}) \mathbf{b}_2$ means proposing a policy on the $\mathbf{b}_2$-ray with the same distance from the origin as $\mathbf{q}$. For status quo policies close to the origin, party 2 is constrained by what the veto player is willing to accept (that is, policies no further away from the origin than
the status quo) and proposes an acceptable policy as close as possible to his bliss point $b_2$.

On the other hand, when $d(q) \geq d(q_2)$, party 2 proposes $\hat{k}_2b_2$, a policy strictly inside the veto player’s acceptance set. Consider, as an example, the status quo $q_1$ in Figure 4. When the status quo is $q_1$, the veto player is willing to accept $b_2$, party 2’s bliss point, because this is just as close to the veto player’s bliss point as $q_1$. This is what a perfectly myopic incumbent would propose. However, a forward-looking party 2 finds it optimal to propose policy $\hat{k}_2b_2$, which is closer to $b_v$ than $b_2$. We call this behavior moderation.

The incentive to moderate is purely strategic and arises from the dynamic nature of the game. By proposing $\hat{k}_2b_2$ rather than $b_2$, party 2 loses in terms of current utility: the policy implemented at the end of this period will be further away from his bliss point. At the same time, however, party 2 gains in terms of expected future utility. Proposing $\hat{k}_2b_2$ rather than $b_2$ means a more favorable status quo in the following period: if the opponent is in power tomorrow, the implemented policy will be $\hat{k}_2b_1$ rather than $b_1$. By moderating, party 2 moves the status quo policy closer to the origin, making the veto player more demanding in requiring policies closer to her bliss point, $b_v$. When party 1 is in power, it is forced to implement a policy closer to $b_v$, which will also be closer to $b_2$ (with respect to what party 1 would be able to implement if a less moderate status quo is in place).

The extent of moderation, that is, the exact location of $\hat{k}_2b_2$, is then determined by the strength of two forces. The first force pushes party 2 in the direction of its bliss point, in an attempt to increase its current utility. The second, strategic force, pushes party 2 in the direction of the veto player’s bliss point, in an attempt to constraint the future behavior of party 1.

We now discuss the equilibrium strategy of party 1. Party 1’s behavior is very similar to party 2’s behavior, with the exception that party 1 does not moderate. This is in spite of the fact that party 1 has a similar strategic incentive to moderate: by moderating, party 1 can constraint the future policies implemented by his opponent, party 2, when in power. However, while all moderate policies reduce party 1’s current utility, not every moderate policy gives party 1 an advantage in terms of future utilities. In order to constrain party 2, party 1 has to moderate to $\hat{k}_2b_1$ or to a policy even closer to the origin, for example, referring again to Figure 4, to $k''b_1$. By moderating less, for example to $k'b_1$, party 1 does not affect the future policies implemented by party 2. Proposing $b_1$,
or $k' b_1$, or $\hat{k}_2 b_1$, or any policy in between, implies that party 2 will propose policy $\hat{k}_2 b_2$ if it grabs power in the following period. However, moderating to $\hat{k}_2 b_1$ or more is too costly for party 1 in terms of foregone current utility. This is because moderation is a strategic substitute and the incentive of party 1 to moderate is weaker than the incentive of party 2 to moderate: party 1 is more likely to be in power, so the gain in future utility he gets from constraining party 2’s future actions is less likely to materialize. In other words, the strategic force pushing party 1’s policies towards $b_v$ is weaker. Since the only degrees of moderation that affect his future utility are too costly (in terms of foregone current utility), party 1 abandons the idea of moderation altogether and, for status quo policies far enough from the origin, proposes his ideal policy, $b_1$.\footnote{Notice that an equilibrium in simple strategies has to be asymmetric also when both parities have the same incumbency advantage, that is, when $r_1 = r_2$. Even if the two parties have the same incentive to moderate, moderation is still a strategic substitute and, provided one party moderates, the opponent has no incentive to moderate at all. As specified by Proposition 1, in this case we have two asymmetric equilibria, one with party 1 moderating and another one with party 2 moderating.}

**Long-Run Policies and Convergence Dynamics.** Proposition 1 also specifies the long run policies that we converge to as a consequence of equilibrium strategies. We have three cases to consider. First, assume the initial status quo, $q^0$, is at least as close to the origin as the strategically induced bliss point of party 2, $\hat{k}_2 b_2$, that is $k(q^0) \leq \hat{k}_2$. In this case, the veto player’s acceptance set is binding for all proposers and all policies will lie at the same distance from the origin as the initial status quo. The policies implemented will be $k(q^0)b_2$ and $k(q^0)b_1$, depending on the identity of the proposer.

Second, assume the initial status quo, $q^0$, is further away from the origin than $\hat{k}_1 b_1$, that is $k(q^0) \geq \hat{k}_1$. If party 1 is the incumbent in the first period, he implements his ideal policy with the consent of the veto player and he does so as long as he stays in power (referring again to Figure 4, this is represented by the arrow labelled as $\phi$). However, as soon as the proposal power shifts to party 2, the policy implemented will move to $\hat{k}_2 b_2$, a policy strictly preferred to the status quo by the veto player (the arrow labelled as $\tau$ in Figure 4). In all future periods, the policy implemented will be at the same distance from the origin, being either $\hat{k}_2 b_2$ or $\hat{k}_2 b_1$, depending on the identity of the incumbent. The third case—$\hat{k}_2 < k(q^0) < \hat{k}_1$—is similar to the second one except that party 1, when in power in the initial period, implements $k(q^0)b_1$ and continues to do so until party 2 comes to power and moves the status quo to $\hat{k}_2 b_2$.\footnote{Notice that an equilibrium in simple strategies has to be asymmetric also when both parities have the same incumbency advantage, that is, when $r_1 = r_2$. Even if the two parties have the same incentive to moderate, moderation is still a strategic substitute and, provided one party moderates, the opponent has no incentive to moderate at all. As specified by Proposition 1, in this case we have two asymmetric equilibria, one with party 1 moderating and another one with party 2 moderating.}
4 Veto Player’s Welfare

The welfare of the veto player can be measured from an ex-ante perspective, that is, at the beginning of the game, by \( V^\theta_v(q^0|\sigma) \), a function of \( r \in \{r_1, r_2\} \) (the probability distribution over the incumbent in the initial period), and of the initial status quo, \( q^0 \). Instead of making arbitrary assumptions on \( \theta \) and \( q^0 \), we assume instead that \( v \) believes \( \theta = 1 \) with probability \( r \in (0,1) \) and that \( q^0 \) is distributed according to a continuous cumulative distribution function \( F(q^0) \) with strictly positive density on \( X \).

**Definition 4.** The ex-ante veto player’s welfare from the dynamic policy competition game is given by:

\[
W(\sigma) = \int_X r V^1_v(z|\sigma) + (1 - r) V^2_v(z|\sigma) d(F(z)).
\]

(7)

When \( \sigma \) describes simple proposing strategies for the incumbent (with associated thresholds \( \hat{k}_1 \) and \( \hat{k}_2 \)) and voting strategy for the veto player of the form \( \sigma_v(p, q) = \text{yes if and only if } d(p) \leq d(q) \), we can express \( W(\sigma) \) as \( W(\hat{k}_1, \hat{k}_2, b) \).

We define \( W \) by explicitly taking into account the effect of \( \hat{k}_i \) and \( b \) because of our focus on the unique SSMPE, which can be fully described by the \( \hat{k}_i \) and \( b \). Lemma 1 shows that the veto player is strictly worse off as the degree of policy moderation observed in equilibrium decreases and as the parties bliss points become more extreme with respect to the veto player’s bliss point.

**Lemma 1.** \( W(\hat{k}_1, \hat{k}_2, b) \) is decreasing in all of its arguments.

**Proof.** See Appendix A1.

5 Comparative Static on Policy Moderation and Volatility

The following proposition formalizes the marginal impact of the model parameters on the strategic force pushing the parties towards moderation:

**Proposition 2** (SSMPE Comparative Static).
For \( \hat{k}_2 = \max \left[ 0, \frac{1-\delta r_1+\delta(1-r_2)\cos \alpha}{1-\delta r_1+\delta(1-r_2)} \right] \) as in the SSMPE from Proposition 1,

\[
\frac{\partial \hat{k}_2}{\partial \alpha} \leq 0 \\
\frac{\partial \hat{k}_2}{\partial \delta} \leq 0 \\
\frac{\partial \hat{k}_2}{\partial r_1} \leq 0 \\
\frac{\partial \hat{k}_2}{\partial r_2} \geq 0
\]  

(8)

Proof. Immediate.

All the comparative static properties are intuitive. The strategic force pushing the weaker party (party 2, given the assumption in Proposition 1) towards moderation gains strength as the conflict between the parties becomes more pronounced (higher \( \alpha \)), as the future becomes more important (higher \( \delta \)), as the probability of the weaker party’s opponent grabbing power increases (higher \( r_1 \)); and it loses strength as the probability the weaker party stays in power for successive periods increases (higher \( r_2 \)).

Combining the comparative static results from Proposition 2 with Lemma 1, we get the following corollary about the marginal impact of antagonism, \( \alpha \), and extremism, \( b \), on the veto player’s welfare.

**Corollary 1** (SSMPE Veto Player’s Welfare Comparative Static). *In the SSMPE characterized in Proposition 1, the veto player’s ex-ante welfare, \( W(\hat{k}_1, \hat{k}_2, b) \),

1. is non-decreasing in \( \alpha \); increasing if \( \hat{k}_2 \neq 0 \), \( \delta > 0 \) and \( r_2 < 1 \);
2. is non-increasing in \( b \); decreasing if \( \hat{k}_2 \neq 0 \).

We illustrate these two results with the help of Figure 5, which shows what happens to the equilibrium proposal strategies when we increase the value of \( \alpha \) (Figure 5a) and \( b \) (Figure 5b), for specific values of \( \delta \), \( r_1 \) and \( r_2 \). We focus here on the case where the initial status quo is sufficiently far from the origin to generate interesting policy dynamics. \(^{10}\) We know from Proposition 1 and the discussion above that, in this case, the equilibrium policies converge to alternation between \( \hat{k}_2 b_2 \) and \( \hat{k}_2 b_1 \). Figure 5 shows the policy implemented whenever party 2 has the power to set the agenda, \( \hat{k}_2 b_2 \). \(^{11}\) While

\(^{10}\) When the initial status quo is closer to the origin than \( k_2 b_2 \), neither \( \alpha \) or \( b \) matter for the degree of moderation of implemented policies. Notice that the size of the set of initial status quo policies for which this is true increases in \( b \) and decreases in \( \alpha \).

\(^{11}\) We do not show the policy implemented in the long run when party 1 sets the agenda, \( \hat{k}_2 b_1 \), in order not to clutter the figures. This policy is very easy to locate: it lies on the horizontal axis, at the same distance from the origin as \( k_2 b_2 \).
this neglects a component of the veto power’s ex-ante welfare – which depends on the whole path of policies, not just on the long-run policies – it captures the main intuition.\footnote{The only other policy that can be observed in equilibrium is b_1, the policy party 1 implements if he sets the agenda in the first period. While α has no impact on this transient policy, the larger is b, the further away is this transient policy from the origin and, in turn, the worse off is the veto player.}
The dashed line in both figures traces \( \hat{k}_2 b_2 \) as \( α \) increases from 0 to \( π \), and \( b \) increases from 0 to 4.

Figure 5: \( \hat{k}_2 b_2 \) as a function of \( α \) and \( b \)

\[ \delta = 1 \text{ and } r_1 = r_2 = 1/2 \text{ or } r_1 = r_2 = 0 \]

(a) effect of \( α \) for \( b = 2 \)

(b) effect of \( b \) for \( α = \frac{1}{2}π \)

We focus first on the impact of antagonism, \( α \) (Figure 5a). When \( α \) increases, the extent of conflict between the two parties increases. This strengthens the strategic force to moderate and the equilibrium proposal of party 2, \( \hat{k}_2 b_2 \), moves closer to the bliss point of the veto player (the origin of the plane). This clearly benefits the veto player since the pair of policies that will be implemented in the long run is closer to \( b_v \). Notice that, when \( \hat{k}_2 = 0 \), a marginal increase in the degree of antagonism does not change the veto player’s welfare. This is because, when \( \hat{k}_2 = 0 \), the long-run equilibrium policies converge to her bliss point, \( b_v \), as soon as party 2 has the power to propose. Nevertheless, the case where \( \hat{k}_2 > 0 \) is not a knife edge occurrence and happens for a wide range of parameters.\footnote{It happens when either \( α, \delta, \) or \( r_1 \) are low enough, or when \( r_2 \) is high enough. Formally, we have \( \hat{k}_2 > 0 \) for \( α \in [0, \frac{π}{2}] \) and any \( δ \); or for \( α \in (\frac{π}{4}, π] \) and \( δ < \frac{1}{\cos{α}} \); or for \( α \in (0, \frac{π}{4}] \), \( δ > \frac{1}{\cos{α}} \); and \( r_1 < \frac{1}{2} + \cos{α} \); or for \( α \in (\frac{π}{4}, π] \), \( δ > \frac{1}{\cos{α}} \), \( r_1 > \frac{1}{2} + \cos{α} \) and \( r_2 > (\frac{1}{2} - r_1 + \cos{α}) \sec{α} \).}

Increasing extremism, \( b \), on the other hand, has the opposite effect (Figure 5b). The long-run equilibrium proposal of party 2, \( \hat{k}_2 \), does not depend on \( b \), which works only through moving the bliss points of the two parties further away from the bliss point of the veto player. The strength of the moderating force does not change with \( b \) and hence \( \hat{k}_2 b_2 \) moves away from the origin when \( b \) increases. Intuitively, increasing extremism
increases both marginal benefits and marginal costs of moderation, leaving their ratio constant. On the other hand, increasing antagonism has no influence on the marginal costs of moderation and increases its marginal benefits. As a result, \( \hat{k}_2 \) depends on \( \alpha \) not on \( b \). And since the veto player’s utility depends on the distance of the implemented policies from her bliss point, scaling up the policies by increasing \( b \) has a negative effect on her welfare.

The next result presents the marginal impact of antagonism and extremism on the variance of the long-run policy outcomes. This is an interesting prediction as it allows us to link preferences’ disagreement to gridlock and policy uncertainty. Denote with \( d_p = d(\hat{k}_2b_2, \hat{k}_2b_1) \) the distance between the two long-run policies in the unique SSMPE from Proposition 1. Along with the recognition probabilities, this measure determines the variance of the long-run equilibrium policies characterized above.

**Proposition 3.** Assume without loss of generality that \( r_1 \geq r_2 \). \( d_p \) is:

1. non-decreasing in \( b \); increasing if \( \alpha > 0 \) and \( \hat{k}_2 > 0 \);
2. increasing in \( \alpha \) when \( \alpha \in [0, \alpha'] \), decreasing in \( \alpha \) when \( \alpha \in [\alpha', \alpha''] \), and constant in \( \alpha \) when \( \alpha \in [\alpha'', \pi] \), where \( \alpha' \in [0, \pi] \), \( \alpha'' \in [0, \pi] \), \( \alpha' < \pi \Leftrightarrow \delta > \frac{1}{1+r_1-r_2+4(1-r_2)} \), \( \alpha'' < \pi \Leftrightarrow \delta > \frac{1}{1+r_1-r_2+4(1-r_2)} \), and \( \alpha' < \alpha'' \) if \( \alpha' < \pi \).

Figure 6: \( d_p \) as a function of \( \alpha \) and \( b \)
\( \delta = 1 \) and \( r_1 = r_2 = 1/2 \) or \( r_1 = r_2 = 0 \)
(a) effect of \( \alpha \) for \( b = 2 \)  
(b) effect of \( b \) for \( \alpha = \frac{1}{2} \pi \)

We illustrate these result with the help of Figure 6, which shows what happens to the long-run policies variance when we increase the value of \( \alpha \) (Figure 6a) and \( b \) (Figure 6b), for specific values of \( \delta \), \( r_1 \) and \( r_2 \). Increasing extremism has a straightforward effect on policy volatility, as it pushes the implemented policies away from the veto player’s ideal point in different directions. Increasing antagonism, on the other hand, has two
effects: on one side, it distances the long-run policies away from each other; on the other, however, it moves them closer to the ideal point of the veto player. The former effects dominates for low $\alpha$, while the latter dominates for high $\alpha$.

6 Extensions and Discussion

Asymmetric Parties’ Bliss Points In the basic model, we assumed that parties are equidistant from the veto player in terms of ideology. Here we explore how the results discussed above change when we relax this assumption and we allow one party to be relatively more extreme than the other. With asymmetric bliss points, a Markovian equilibrium in simple strategies still exists and it is generically unique. This equilibrium will be asymmetric, with one party moderating towards the bliss point of the veto player and the other one implementing his ideal policy when not constrained by the acceptance set of the veto player. The reason for an asymmetric equilibrium is the same we discussed for the case of symmetric parties: moderation is a strategic substitute and only one of the two parties will moderate in equilibrium. When there is a source of asymmetry, the party who moderates in the unique asymmetric equilibrium is the one who benefits the most from moderation. When $r_1 > r_2$ (that is, the incumbency advantage of party 1 is higher) and $d(b_1) = d(b_2)$ (that is, parties are equidistant from the veto player), the party who moderates is party 2, the one more concerned about power switching.

With asymmetric bliss points, a relatively more moderate party has a higher incentive to moderate. This is because moderation imposes a smaller loss in terms of current utility to the party whose bliss point is closer to the origin. Moreover, the extent of optimal moderation is not a function of the bliss points’ exact locations or of the Euclidean distance between the parties’ bliss points, but only of the degree of ideological conflict, as summarized by the angle $\alpha$ (which does not change when a party’s bliss point move further away from the origin along the same ray). This is because the upside of moderation consists in limiting the policies the opponent can implement in the future, through the acceptance set of the veto player. When both parties are constrained by the veto player’s desires, a more extreme party will implement the same policy as a more moderate party (along the ray connecting the origin with his bliss point, making the veto player indifferent between acceptance and rejection). As a consequence, if $r_1 \geq r_2$ and $d(b_1) > d(b_2)$ (that is, party 1 is relatively more extreme than party 2), both asymmetries give a higher incentive to moderate to party 2, and the unique SSMPE is
the one characterized in Proposition 1.\textsuperscript{14}

When the more moderate party has a higher incumbency advantage, then the two sources of asymmetry do not unambiguously determine which party has a higher incentive to moderate and this depends on what asymmetry has a stronger effect. In particular, consider $\kappa_i$ as defined in the proof of Proposition 1. In the asymmetric game we will have an equilibrium in simple strategies characterized by $\hat{k}_i = \min\{\kappa_i, 0\}$ and $\hat{k}_{-i} = 1$ where $i = 2$ if $\kappa_1 b_1 > \kappa_2 b_2$—with $b_i = d(b_i)$—and $i = 1$ otherwise. As in the game with symmetric bliss points, in this equilibrium the long-run outcomes consist of two policies (on the two rays connecting the origin and the two parties’ bliss points) located at a distance of $\min\{\kappa_1 b_1, \kappa_2 b_2, 0\}$ from the origin.\textsuperscript{15}

**Correlation Between Antagonism and Extremism** In the basic model, extremism ($b$) and antagonism ($\alpha$) are uncorrelated and we focus on the effect of increasing one, while keeping the other constant. However, we can think of a plausible political scenario in which antagonism and extremism are correlated: as the degree of conflict between the parties increases, so does the distance between the parties and the (formal or informal) veto player. Under this alternative assumption, the relevant comparative static is the impact of $b(\alpha)$ on the implemented policies and the veto player’s welfare. The long run policies crucially depend on the position of $\hat{k}_2 b_2$, where the equilibrium degree of moderation, $\hat{k}_2$, is weakly decreasing in $\alpha$, and $b_2$ is strictly increasing in $b$. When we increase $\alpha$ and $b$ at the same time, whether the long run policies are more or less extreme, depends on whether the marginal effect of $\alpha$ on moderation prevails over the marginal effect of $b$ on extremism (which will in turn depend on the relative speed at which $b$ and $\alpha$ increase). We can however make some interesting observations. First, note, that—as $\alpha$ goes to $\pi$ and $b$ goes to $\infty$—the long run policies can converge to either the origin or a point infinitely far from the origin. When $\hat{k}_2$ is positive for every $\alpha \in [0, \pi]$, then, as $b$ goes to $\infty$, $\hat{k}_2 b_2$ keeps growing with $b$ (increasing the variance of the long run policies and decreasing the veto player’s welfare). On the other hand, when $\hat{k}_2 = 0$ for some $\alpha < \pi$, then $\hat{k}_2 b_2 = 0$ for a finite $b$, and marginally increasing $b(\alpha)$ beyond the level for which $\hat{k}_2 = 0$ has no impact on the long run policies. Finally, if $\hat{k}_2 = 0$ only for $\alpha = \pi$, whether the moderating effect of $\alpha$ prevails and the long run policies converge to the origin (rather than to $\infty$) depends on the relative speed of convergence of $\alpha$ to $\pi$ and $b$.

\textsuperscript{14} Where the distance $\hat{k}_i$ is now relative to the $d(b_i)$ distance from $b_i$.

\textsuperscript{15} When $d(b_1) < d(b_2)$ and $r_1 > r_2$, the parties’ strategies in a SSMPE can display jumps in the state variable, but the qualitative predictions are unchanged.
to $\infty$ (but notice that this is a knife edged case).

**Legislative Bargaining**  Our model can also be interpreted as a model of dynamic bargaining in a legislature, where the power to set the agenda is in the hand of the majority party and each party maximizes the utility of the party median legislator or the party leaders (as in the procedural cartel theory introduced by Cox and McCubbins 1993, 2005). To view our model as a legislative bargaining model, reinterpret the veto player as the median legislator in the assembly. In this sense, our paper is closely related to the dynamic legislative bargaining model of Baron (1996). There are three crucial differences between our paper and Baron (1996). First, the policy space in our paper is two-dimensional, rather than unidimensional, with an institutional veto player (rather than the median legislator on the floor) being decisive for the passage of the policy proposals. Second, in our paper only two parties or party leaders (rather than any legislator in the assembly) have the power to set the agenda. In particular, the median legislator, whose approval is decisive for passage of a proposal, has never the power to propose nor is represented by a party that shares its preferences over policies. Third, our proposer recognition rule is more general and allows for inertia in the identity of the proposer, rather than being an independent draw in each round.

**Direct Democracy**  Our model can also be interpreted as a model of direct democracy in which the incumbent party can change the status quo policy but the citizens can call a referendum to revoke any legislation they do not approve of (as in Romer and Rosenthal 1979, Matsusaka and McCarty 2001, and Gerber 1996); or a model where citizens cannot repeal undesired legislation with a referendum but can exercise oversight over their representative and have the option to take the streets and riot when the government tries to implement a policy that they do not like, relative to the status quo (with this action likely to obtain the desired result). In this case, the institutional veto player is replaced by a median or decisive voter and the welfare statements on the marginal effect of elite antagonism and extremism pertain to the electorate as a whole, rather than to a particular veto player.

**Endogenous Re-Election Probabilities**  In the basic model, the Markov process governing the power to set the agenda is exogenous and the policy proposed by the incumbent does not influence its probability of being re-elected (or maintaining the majority in the legislative assembly). An alternative modeling assumption is that, when
the incumbent proposes a more moderate policy, the electorate rewards him with a higher probability of re-election. While such an extension is beyond the scope of this paper, we can note that, in this case, there would be an additional incentive to implement a policy closer to the veto player’s (or the representative voter’s) ideal point. As a consequence, the competition between ideological agenda setters would get even more intense and the conclusion that antagonism increases the leverage of the veto player (or the representative voter) over policies would still hold and actually be reinforced.

**Antagonism and Extremism in U.S. Congress**  One of the main contributions of the paper is to propose two novel measures of ideological disagreement among political actors, which have different but complementary effects on policy outcomes. It is interesting to understand how our measures relate to existing notions of ideological conflict and how they can expand our understanding of polarization and its consequences. A large body of research has been devoted to identifying ideological polarization in the United States. McCarty, Poole and Rosenthal (1997, 2006) use roll call voting to estimate legislators’ ideal points on a bi-dimensional space and measure the level of political polarization in Congress with the difference between the Republican and Democratic Party means on the first dimension. By this measure, polarization is now at a post-Reconstruction high in both the House and Senate. They also show that the increase in party polarization has reduced the dimensionality of political conflict: while in the past parties divided internally on a variety of issues—especially those related to race and other regional issues—the importance of the second dimension in explaining roll-call votes has waned and the party difference in second dimension means has converged over time.

In terms of our model, an increase in ideological conflict can reflect an increase in antagonism, $\alpha$, or an increase in extremism, $b$, with markedly different consequences on policy outcomes. To understand what aspect of polarization has changed over time, we rely on two observations. First, we note that, an increase in $\alpha$ is associated with an increased distance between the ideal points of the two parties on both dimensions when $\alpha \in [0, \pi/2]$ and an increased distance on one dimension, together with a reduced distance on the other, when $\alpha \in [\pi/2, \pi]$. Competition is summarized by a single dimension only when $\alpha$ is at its maximum. Second, if we keep $\alpha$ fixed and raise $b$, we increase the distance between the two parties in both dimensions. This suggests that we can interpret the historical increase of polarization in the U.S. Congress as an increase in the antagonism between the two parties. However, now that antagonism has reached its maximum level
(as evidenced by the reduction of conflict to a single dimension), any further increase
in the observed level of polarization has to be due to an increase in extremism. We
conclude that increased polarization between the two main American parties might have
benefitted the institutional veto players or the moderate voters in its early phase but
that, at this point, any further increase is likely to have the opposite consequences.

7 Conclusions

While many commentators and scholars diagnose a sharp and increasing ideological
divide between the main American parties, both the popular press and the existing lit-
erature are somewhat unclear about what exactly constitutes polarization and how one
can measure this concept. In this paper, we study an environment where two ideological
and forward-looking parties alternate in office and the multidimensional policies they
propose have to be approved by a veto player. We use two different measures to describe
the political environment and the degree of conflict among agents’ preferences. The first
measure, which we label extremism, is the ideological distance of each party from the
veto player. The second measure, which we call antagonism, is the ideological distance
that separates the two parties from each other and summarizes the degree of political
competition between the agenda setters. These two measures coincide in a one dimen-
sional policy space, where the ideological distance between the two parties can increase
only as they move further away from the veto player. However, they do not coincide
in a two-dimensional setting: here, the two parties can be very close (when they share
views on both dimensions) or very different (when they are perfectly opposed in one
dimension), without altering their overall distance from the veto player. We show that
a stationary Markov perfect equilibrium of this game exists and we fully characterize
it for any discount factor, initial status quo policy, and any degree of extremism and
antagonism. In this equilibrium, increasing the degree of extremism reduces the welfare
of the veto holder. On the other hand, increasing antagonism increases the influence
exerted by the veto holder on long run policy outcomes.

We stress two possible interpretations of our model. One interpretation sees as veto
player an institutional veto holder, such as, in the context of American politics, the U.S.
President, the median legislator in Congress, the Speaker of the House of Representatives
or the Senate, or the Supreme Court. In the second interpretation, we can think of the
veto player as a representative voter, in the context of a model of direct democracy. In
the remainder, we discuss the empirical implications of our analysis for both scenarios.

Our analysis suggests that institutional veto holders can exert a substantial influence over policy outcomes, even when the initial status quo policy disadvantages them and when they have no (direct or indirect) ability to set the agenda. The degree of this influence positively depends on the degree of ideological conflicts between the parties in the legislative assembly. This result suggests a set of novel and testable empirical implications. An immediate empirical implication of this result is that the observed influence of institutional veto holders over implemented policies should be higher in periods of high antagonism between parties (for example, the current decade in the U.S.) and lower in periods of low antagonism (for example, the 50s and 60s in the U.S.). A second empirical implication is that the allocation of veto powers and the identity of veto holders are more important (in the sense of affecting final policy outcomes) when parties are more antagonized. This, for example, could generate a higher conflict inside each party and between parties over the nomination of key political figures like the speakers of the two Chambers or the members of the Supreme Court. A third prediction delivered by the model is that, even when parties’ ideologies are symmetric, moderation is a strategic substitute and only one party moderates, with the other party sticking to his guns (at least in the initial periods, when the veto holder is disadvantaged by the status quo). This prediction is in line with the asymmetric evolution of observed polarization in the American Congress, where the Republican representatives are relatively more extreme than the Democratic ones. Finally, as pointed by Baker, Bloom and Davis (2013) and Fernandez-Villaverde, Guerron-Quintana, Kuester and Rubio-Ramirez (2011), the last decades have been associated with greater-than-historical economic policy uncertainty and volatility in the United States. The theoretical predictions of our model suggest that this could be linked to ideological conflict, through two different channels: both an increase in extremism and an increase in antagonism (starting from a low level) can lead to higher volatility of policy outcomes. Interestingly, the period between the 1970s and 2011 has also been a period of higher-than-historical political polarization, as highlighted by McCarty, Poole and Rosenthal (2006) on the basis of legislators’ ideal points estimated from roll call votes.

In some instances, the institutional veto holders are chosen to represent the preferences of the electorate at large. Moreover, our model can easily be interpreted as a model of direct democracy where incumbent parties can change the status quo only with the consent of a moderate, representative voter. When we read our model using these lenses, does it suggest that the American electorate is better off with more polarized parties?
The answer is likely to depend on the dimensionality of the political conflict. If the political competition is on one dimension—as argued by the theory of “conflict displacement” (Sundquist 1983, Carmines and Stimson 1989, Miller and Schofield 2003) and suggested by the roll call analysis of McCarty, Poole and Rosenthal (2006) for the current historical period—then increased elite polarization coincides with increased extremism, using the language of our model, and this is likely bad news for the moderate voters: as polarization increases, both the distance between parties and the distance of both parties from the moderate voters increase, and the observed policies will be more extreme (even when parties are forward looking and policy motivated). If instead the political competition and the increased ideological conflict is on multiple, correlated dimension—as argued by the theory of “conflict extension” (Layman and Carsey 2002a)—this might be good news for moderate voters: the heightened competition between parties will not favor the civility of the political discourse and might lead to policy gridlock, but forward looking and policy motivated parties will propose more moderate policies.

References


A1 Proofs

Proof of Proposition 1

Throughout the proof, define $\kappa_i$ as:

$$\kappa_i = \frac{1 - \delta r_{-i} + \delta (1 - r_i) \cos \alpha}{1 - \delta r_{-i} + \delta (1 - r_i)}$$

Let $\hat{k}_1$ and $\hat{k}_2$ be the parameters associated with simple strategies and let $\hat{k}_1^*$ and $\hat{k}_2^*$ be the parameters associated with the simple strategies of a SSMPE (as described in Definition 3). When talking about both parameters jointly, we will use $\hat{k} = (\hat{k}_1, \hat{k}_2)$ and $\hat{k}^* = (\hat{k}_1^*, \hat{k}_2^*)$. It is easy to confirm, given $r_i \in [0,1]$ for $i \in \{1,2\}$, $\delta \in [0,1)$ and $\alpha \in [0,\pi]$, that $\kappa_1 \leq 0$ and $\kappa_2 \leq 0$ cannot hold simultaneously, that $\kappa_i \leq 1$ for $i \in \{1,2\}$ and that $\kappa_i \geq \kappa_{-i} \Leftrightarrow r_1 \geq r_{-i}$.

First notice that, when parties use simple strategies characterized by $\hat{k}$, for any strategy of the veto player, $\sigma_v$, and two different status quo policies $q'$ and $q''$ with $d(q') = d(q'')$, $V_v^j(q'|(\hat{k},\sigma_v)) = V_v^j(q''|(\hat{k},\sigma_v))$ for $j \in \{1,2\}$. It follows that, in a SSMPE, $\sigma_v$ has to satisfy $\sigma_v(p,q') = \sigma_v(p,q'')$ for any $(p,q',q'') \in X^3$ with $d(q') = d(q'')$, since the stage utility of the veto player $u_v(q) = -d^2(q)$. Furthermore, when parties use simple strategies characterized by $\hat{k}$, the continuation value of the veto player, $V_v^j(q|(\hat{k},\sigma_v))$, is non-increasing in $d(q)$ for $j \in \{1,2\}$. Since $u_v(q)$ is decreasing in $d(q)$, in a SSMPE, $\sigma_v$ has to satisfy $\sigma_v(p,q) = \text{yes}$ if and only if $d(p) \leq d(q)$.$^{16}$ This proves part three of the proposition. Because of the simplicity of the veto player’s behavior in any SSMPE, we suppress $\sigma_v$ from the notation below.

The set of policies the veto player accepts given the status quo $q$ (that is, her acceptance set) is a (closed) circle centered at $b_v = (0,0)$, with radius $d(q)$. This, together with the fact that parties use simple strategies characterized by $\hat{k}$, and that $\hat{k}_1 \geq \hat{k}_2$ (something we show below) proves part four of the proposition. In particular, when $k(q) \leq \hat{k}_2$, party $i \in \{1,2\}$ proposes $k(q)b_i$ whenever in power. When $k(q) \geq \hat{k}_2$, the first time party 2 sets the agenda, it proposes $\hat{k}_2b_2$ and, from then on, the policies alternate between $\hat{k}_2b_1$ and $\hat{k}_2b_2$, depending on the identity of the proposer. For $r_1 = 1$, if party 1 is in power in the initial period, it stays in power indefinitely, and the

$^{16}$ Notice that this does not depend on the veto player’s discount factor $\delta_v$. The reason is that the veto player’s stage utility $u_v$ and her continuation value $V_v^j$ have the same shape. No matter how she weights the present and the future, her preferences over pairs of policies remain the same.
The continuity is easy to see as the simple strategies characterized by \( \hat{k} \) give rise to an expected value function, \( V_i(q|\hat{k}) \), that is a continuous function of \( q \). For the convergence specified in the fourth part of the proposition need not obtain. When this happens, the path of policies is either \( k(q)b_1 \) or \( \hat{k}_1b_1 \) depending on whether \( k(q) \leq \hat{k}_1 \). This is part five of the proposition, which already uses SSMPE \( \hat{k}_1 = 1 \).

The stage utility \( i \in \{1, 2\} \) derives from \( k b_j \) where \( k \geq 0 \) and \( j \in \{1, 2\} \) is \( u_i(k b_j) = -d^2(k b_j, b_i) \). This is a continuous and differentiable function of \( k \). We have:

\[
\frac{\partial (-d^2(k b_i, b_i))}{\partial k} = -2(k d^2(b_{-i}) - b_{-i} \cdot b_i) = -2b^2(k - \cos \alpha)
\]

\[
\frac{\partial (-d^2(k b_i, b_i))}{\partial k} = -2(k d^2(b_i) - b_i \cdot b_i) = -2b^2(k - 1)
\]

\[
\frac{\partial^2 (-d^2(k b_j, b_i))}{\partial^2 k} = -2d^2(b_j) = -2b^2
\]

This proves that \( u_i(k b_j) \) is a concave function of \( k \).

Consider a status quo \( q \) and a given \( \hat{k} \). When party \( i \) is in power, it proposes policy \( p^* \in \arg \max_{d(x) \leq d(q)} u_i(x) + \delta V_i(x|\hat{k}) \). For any \( \hat{k} \) and any two policies \( p \in X \) and \( p' \in X \) with \( d(p) = d(p') \), we have \( V_i(p|\hat{k}) = V_i(p'|\hat{k}) \), so that \( p^* \) has to lie on the ray starting at \( b_v = (0, 0) \) and passing through \( b_i \), and hence can be written as \( p^* = k b_i \) for some \( k \geq 0 \). Denote by \( U_i(k|\hat{k}) = u_i(k b_i) + \delta V_i(k b_i|\hat{k}) \) the expected utility of party \( i \in \{1, 2\} \) from proposing policy \( k b_i \) when the parties use simple strategies characterized by \( \hat{k} \). The following lemma shows key properties of \( U_i(k|\hat{k}) \).

**Lemma A1.** Fix \( \hat{k}_1 \geq 0, \hat{k}_2 \geq 0 \) and \( i \in \{1, 2\} \). \( U_i(k|\hat{k}) \), as a function of \( k \geq 0 \), is continuous, differentiable except when \( k = \hat{k}_1 \) or \( k = \hat{k}_2 \), strictly concave on each interval on which it is differentiable and \( \frac{\partial U_i(k|\hat{k})}{\partial k} > 0 \) for \( k < c \), \( \frac{\partial U_i(k|\hat{k})}{\partial k} = 0 \) for \( k = c \), and \( \frac{\partial U_i(k|\hat{k})}{\partial k} < 0 \) for \( k > c \) where

\[
c = \kappa_i \quad \text{if} \quad k < \min\{\hat{k}_1, \hat{k}_2\}
\]

\[
c = \kappa_i \quad \text{if} \quad \hat{k}_i < k < \hat{k}_{-i}
\]

\[
c = 1 \quad \text{if} \quad \hat{k}_{-i} < k < \hat{k}_i
\]

\[
c = 1 \quad \text{if} \quad \max\{\hat{k}_1, \hat{k}_2\} < k
\]

**Proof.** The continuity is easy to see as the simple strategies characterized by \( \hat{k} \) give rise to an expected value function, \( V_i(q|\hat{k}) \), that is a continuous function of \( q \). For the
remaining properties, we have to derive \( V_i^t(q|\hat{k}) \) explicitly. It clearly satisfies

\[
V_i^t(q|\hat{k}) = r_i \left[ -d^2(z_i, b_i) + \delta V_i^t(z_i|\hat{k}) \right] \\
\quad + (1 - r_i) \left[ -d^2(z_{-i}, b_i) + \delta V_i^{-i}(z_{-i}|\hat{k}) \right]
\]

\[ V_i^{-i}(q|\hat{k}) = (1 - r_{-i}) \left[ -d^2(z_{-i}, b_i) + \delta V_i^{-i}(z_{-i}|\hat{k}) \right] \\
\quad + r_{-i} \left[ -d^2(z_{-i}, b_i) + \delta V_i^{-i}(z_{-i}|\hat{k}) \right]
\]

where \( z_i \) and \( z_{-i} \) depend on \( k(q) \)

\[
\begin{align*}
z_i &= k(q)b_i \\
z_{-i} &= k(q)b_{-i} \quad \text{if} \quad k(q) < \min \{\hat{k}_1, \hat{k}_2\} \\
z_i &= \hat{k}_ib_i \\
z_{-i} &= k_{-i}b_{-i} \quad \text{if} \quad \hat{k}_i < k(q) < \hat{k}_{-i} \\
z_i &= k(q)b_i \\
z_{-i} &= \hat{k}_{-i}b_{-i} \quad \text{if} \quad \hat{k}_{-i} < k(q) < \hat{k}_i \\
z_i &= \hat{k}_ib_i \\
z_{-i} &= k_{-i}b_{-i} \quad \text{if} \quad \max \{\hat{k}_1, \hat{k}_2\} < k(q).
\end{align*}
\]

Straightforward algebra, using \( k(k^\prime b_i) = k(k^\prime b_{-i}) = k^\prime \) along with \( V_i^t(q|\hat{k}) = V_i^t(k(q)b_i|\hat{k}) \) and \( V_i^{-i}(q|\hat{k}) = V_i^{-i}(k(q)b_{-i}|\hat{k}) \), then shows, using the same order of cases as in (A2) and summarizing with \( \chi \) all the terms constant in \( k \):

\[
\begin{align*}
U_i(k|\hat{k}) &= -\frac{d^2(kb_i, b_i)[1 - \delta r_{-i}] - d^2(kb_{-i}, b_i)\delta [1 - r_i]}{(1 - \delta)(1 + \delta(1 - r_{-i} - r_i))} \\
U_i(k|\hat{k}) &= -\frac{d^2(kb_i, b_i)[1 - \delta r_{-i}] - d^2(kb_{-i}, b_i)\delta [1 - r_i]}{1 - \delta r_i} + \chi
\end{align*}
\]

Direct verification then shows all the remaining properties of \( U_i(k|\hat{k}) \). \( \square \)

Using \( U_i(k|\hat{k}) \) we can rewrite the optimization problem of party \( i \in \{1, 2\} \) regarding which policy to propose, when in power with status quo \( q \), as \( \max_{0 \leq k \leq k(q)} U_i(k|\hat{k}) \). To find a SSMPE, we need to find a \( \hat{k}^* \) such that the solution to this optimization problem under \( \hat{k}^* \) can be described by \( \hat{k}^* \).

First notice that for any \( \hat{k} \), \( U_i(k|\hat{k}) \) is decreasing in \( k \) for \( k > 1 \). Hence \( \hat{k}_1^* \in [0,1] \) and \( \hat{k}_2^* \in [0,1] \). We now show that \( \hat{k}_1^* < 1 \) and \( \hat{k}_2^* < 1 \) cannot characterize an SSMPE. Assume, towards a contradiction, that \( \hat{k}_1^* < 1 \) and \( \hat{k}_2^* < 1 \) and denote by \( \hat{k}_1^* = \max \{\hat{k}_1^*, \hat{k}_2^*\} \)
(in case of equality the choice is arbitrary). As \( \hat{k}_m^* < 1 \), by lemma A1, \( U_m(k|\hat{k}^*) \) is increasing in \( k \) on \( (\hat{k}_m^*, 1) \), yet \( m \) proposes \( \hat{k}_m^* b_m \) for any \( q \) such that \( k(q) \geq \hat{k}_m^* \), a contradiction to \( m \) maximizing \( U_i(k|\hat{k}^*) \).

This leaves three possible cases, \( \hat{k}_1^* = 1 \) with \( \hat{k}_2^* < 1 \), \( \hat{k}_1^* < 1 \) with \( \hat{k}_2^* = 1 \) and \( \hat{k}_1^* = 1 \) with \( \hat{k}_2^* = 1 \). By lemma A1, the last case is clearly only possible when \( \delta = 0 \) or when \( \alpha = 0 \) or when \( r_2 = 1 \), in which case the former two cases cannot characterize an SSMPE. For \( \delta = 0 \) or \( \alpha = 0 \) or \( r_2 = 1 \), \( \hat{k}_1^* = 1 \) and \( \hat{k}_2^* = \max[0, \kappa_2] \).

To characterize the remaining cases, assume \( \delta \in (0, 1) \) and \( \alpha \in (0, \pi] \). When \( r_1 > r_2 \), we have \( \kappa_1 > \kappa_2 \), \( \hat{k}_1^* < 1 \) and \( \hat{k}_2^* = 1 \) cannot characterize an SSMPE by lemma A1, so we must have \( \hat{k}_1^* = 1 \) and \( \hat{k}_2^* = \max[0, \kappa_2] \). When \( r_1 = r_2 \), we have \( \kappa_1 = \kappa_2 \) and both \( \hat{k}_1^* = 1 \) with \( \hat{k}_2^* = \max[0, \kappa_2] \) and \( \hat{k}_1^* = \max[0, \kappa_1] \) with \( \hat{k}_2^* = 1 \) characterize SSMPE. □

**Proof of Lemma 1**

Let \( \sigma = (\hat{k}_1, \hat{k}_2, \sigma_v) \) describe simple (not necessarily SSMPE) strategies for the two parties and strategy for the veto player of the form \( \sigma_v(p, q) = \text{yes} \text{ if and only if } d(p) \leq d(q) \). \( V_v^i(q|\sigma) \) in the definition of \( W(\sigma) \) is the expected utility of the veto player from (random) path of policies generated by play according to \( \sigma \), starting from the status quo \( q \) with \( r_i \) determining the party in power in the initial period.

Suppressing the dependence on \( \sigma_v \), denote by \( P_{\hat{k}_1, \hat{k}_2, b, q, i} = \{p_0, p_1, \ldots\} \) fully deterministic path of policies generated by simple strategies \( \hat{k}_1, \hat{k}_2 \) for given path of the two parties alternating in power, in a model with \( b \), starting from status quo \( q \) and \( i \in \{1, 2\} \) in power. Then

\[
V_v^i(q|\sigma) = E \left[ r_i U_v \left( P_{\hat{k}_1, \hat{k}_2, b, q, i} \right) + (1 - r_i) U_v \left( P_{\hat{k}_1, \hat{k}_2, b, q, -i} \right) \right]
\]  

(A6)

where the expectation is over all possible paths of the two parties alternating in power. It is easy to confirm that for each \( p_i \) in \( P_{\hat{k}_1, \hat{k}_2, b, q, i} \), \( d(p_i) \) is non-decreasing in \( \hat{k}_1, \hat{k}_2 \) and \( b \) and hence \( V_v^i(q|\sigma) \) is non-increasing in the same parameters for any \( q \in X \) and \( i \in \{1, 2\} \). Additionally, there will always be a set of status quo policies in \( X \) with positive measure, such that the dependence will be strict, proving the lemma. □
Proof of Proposition 3

By proposition 1, \( \hat{k}_1 = 1 \) and \( \hat{k}_2 = \max \{0, \kappa_2\} \), where \( \kappa_2 = \frac{1-\delta r_1 + \delta (1-r_2) \cos \alpha}{1-\delta (1-r_2)} \). To economize on notation, denote \( \kappa_n = 1-\delta r_1 + \delta (1-r_2) \cos \alpha \) and \( \kappa_d = 1-\delta r_1 + \delta (1-r_2) \), so that \( \kappa_2 = \frac{\kappa_n}{\kappa_d} \). Note that, given the assumptions of the model, \( \kappa_d > 0 \).

To show the first part, using polar to Cartesian coordinates transformation \( \hat{k}_2 \mathbf{b}_2 = (\hat{k}_2 b \cos \alpha, \hat{k}_2 b \sin \alpha) \), \( d_p \) expressed as \( d_p = 2b \hat{k}_2 \sin \frac{\alpha}{2} \), so that \( d_p \) is clearly non-decreasing in \( b \). It is increasing if \( \alpha > 0 \) and \( \hat{k}_2 > 0 \).

To show the second part, we first note that \( d_p = 0 \) when \( \hat{k}_2 = 0 \). We also know that \( \hat{k}_2 \) is non-increasing in \( \alpha \), so that when \( \hat{k}_2 = 0 \) for some \( \alpha'' \), then \( \hat{k}_2 = 0 \) for \( \forall \alpha \geq \alpha'' \).

Smallest such \( \alpha'' \) is easily found by setting \( \kappa_n = 0 \), or \( \alpha'' = \arccos \left[ -\frac{1-\delta r_1}{\delta (1-r_2)} \right] \), which is well defined only if \( \delta > 0 \), \( r_2 < 1 \) and \( -1 \leq -\frac{1-\delta r_1}{\delta (1-r_2)} \leq 1 \). The last condition, from \( \{x | \arccos x \in \mathbb{R}\} = [-1, 1] \), implies

\[
\delta \geq \frac{1}{1+r_1-r_2}. \tag{A7}
\]

Note also that \( \delta \geq \frac{1}{1+r_1-r_2} \) implies \( \delta > 0 \) and \( r_2 < 1 \); when \( \delta = 0 \) the inequality cannot hold as \( \frac{1}{1+r_1-r_2} > 0 \) and \( r_2 = 1 \) implies \( r_1 = 1 \) and hence \( \delta \geq 1 \), which is impossible. Setting \( \alpha'' = \pi \) when (A7) fails and noting that range or arccos is \([0, \pi]\), we have \( \alpha'' \in [0, \pi] \). From \( \arccos x = \pi \Leftrightarrow x = -1 \) we also have \( \alpha'' < \pi \Leftrightarrow \delta > \frac{1}{1+r_1-r_2} \).

It turns out to be algebraically more convenient to work with \( d_p^2 \), which is permissible as \( d_p \) denotes distance. We have, when \( \hat{k}_2 \neq 0 \),

\[
\frac{\partial d_p^2}{\partial \alpha} = \frac{\sin \alpha}{\kappa_d^2} \kappa_n \left[ \kappa_d + 3 \delta (1-r_2) (\cos \alpha - 1) \right]. \tag{A8}
\]

If the derivative is zero for some \( \alpha' \in (0, \pi) \), then it is negative for \( \forall \alpha \in (\alpha', \pi) \), if \( \hat{k}_2 \neq 0 \) still holds, and positive for \( \forall \alpha \in (0, \alpha') \). This follows because \( \hat{k}_2 \neq 0 \), \( \alpha' \in (0, \pi) \) and the derivative equal to zero implies that the term in the square brackets is zero and it is clearly decreasing in \( \alpha \) for \( \alpha \in (0, \pi) \) (\( \delta = 0 \) and \( r_2 = 1 \) are ruled out by the square bracketed terms being zero). The bracketed terms is zero if \( \alpha = \alpha' \) where \( \alpha' = \arccos \left[ 1 - \frac{\kappa_d}{3 \delta (1-r_2)} \right] \). \( \alpha' \) is well defined if \( \delta > 0 \), \( r_2 < 1 \) and \( -1 \leq 1 - \frac{\kappa_d}{3 \delta (1-r_2)} \leq 1 \). The last condition implies

\[
\delta \geq \frac{1}{1+r_1-r_2 + 4(1-r_2)} \tag{A9}
\]
and, using similar argument as for $\alpha''$, implies $\delta > 0$ and $r_2 < 1$. Setting $\alpha' = \pi$ when (A9) fails, we thus have $\alpha' \in [0, \pi]$ and $\alpha' < \pi \iff \delta > \frac{1}{1+r_1-r_2+4(1-r_2)}$.

To show that $\alpha' < \alpha''$ if $\alpha' < \pi$, from definition of $\alpha'$ and $\alpha''$ and the fact that $\arccos$ is decreasing function we need to show

$$- \frac{1 - \delta r_1}{\delta (1 - r_2)} < 1 - \frac{\kappa_d}{3\delta (1 - r_2)}$$

which, because $\alpha' < \pi$ implies $\delta > 0$ and $r_2 < 1$, is easily checked to hold.

Last thing we need to show is that $\alpha'$ and $\alpha''$ determine different intervals on which $d_p$ is increasing, decreasing and constant in $\alpha$. 

**Case 1:** When $\delta \geq \frac{1}{1+r_1-r_2}$, that is when (A7) holds, then $\hat{k}_2 = 0$ for $\forall \alpha \in [\alpha'', \pi]$ where $\alpha'' \in [0, \pi]$ and $d_p$ is constant in $\alpha$ on the same interval. (A7) implies that (A9) holds strictly, so there exists $\alpha' \in (0, \pi)$ with $\alpha' < \alpha''$ such that $d_p$ is increasing $\forall \alpha \in [0, \alpha']$ and decreasing $\forall \alpha \in [\alpha', \alpha'']$. 

**Case 2:** When $\delta < \frac{1}{1+r_1-r_2}$ but (A9) holds, then $\hat{k}_2 \neq 0$ for $\forall \alpha \in [0, \pi]$ and there exists $\alpha' \in [0, \pi]$ such that $d_p$ is increasing $\forall \alpha \in [0, \alpha']$ and decreasing $\forall \alpha \in [\alpha', \pi]$ respectively. 

**Case 3:** When (A9) fails then $d_p$ is increasing $\forall \alpha \in [0, \pi]$. 

□

### A2 Model with General Utility

The model analyzed in this section is identical to the one in the main part of the paper except for the stage utility player $i \in \{1, 2, v\}$ derives from policy $p$ which now is:

$$u_i(p) = f(d(p, b_i))$$  \hspace{1cm} (A11)

where $f : [0, \infty) \to \mathbb{R}$ is a continuous, decreasing and concave function in $d(p, b_i)$. We also assume that $f$ is twice continuously differentiable on $[0, \infty)$. Notice that these assumptions allow $f'(0) = 0$ and $f(x) = -x^2$, so that the model in the paper is a special case of the model analyzed here.

For space considerations, we refrain from repeating the arguments leading to Proposition 1 and stress only those aspects of the analysis that differ considerably. An argument similar to the one from the proof of Proposition 1 shows that, since players use simple strategies, it is i) optimal for $i \in \{1, 2\}$ to propose policies located only on the $b_i$-ray and ii) optimal for $v$ to approve the proposed $p$ when the status-quo is $q$ if and only if $d(p) \leq d(q)$.
In a SSMPE, the extent of moderation of party $i \in \{1, 2\}$ is $k$ which maximizes $\tilde{U}_i(k)$ where, for $k \geq 0$,

$$
\tilde{U}_i(k) = u_i(k b_i)(1 - \delta r_{-i}) + u_i(k b_{-i})\delta(1 - r_i) \\
= f(d(k b_i, b_i))(1 - \delta r_{-i}) + f(d(k b_{-i}, b_i))\delta(1 - r_i).
$$

(A12)

$\tilde{U}_i(k)$ is the dynamic utility player $i$ receives, devoid of any constant terms, when she proposes policy $k b_i$ and player $-i$, not moderating to a larger extent, proposes $k b_{-i}$. The derivation of $\tilde{U}_i(k)$ is similar to the derivation of (A5) and is not repeated here.

**Condition 1** (Non trivial model). $\delta \in (0, 1)$, $r_i < 1$ and $\alpha \neq \pi$.

**Condition 2** (Preserving concavity). If $\alpha = \pi$, then $f'' < 0$.

**Lemma A2.** $\tilde{U}_i(k)$ is

1. continuous and twice continuously differentiable for $k \neq 1$
2. increasing for $k \in [0, \cos \alpha]$
3. decreasing for $k \geq 1$
4. if conditions 1 and 2 hold, strictly concave for $k \in (\max \{0, \cos \alpha\}, 1)$
5. if condition 1 fails, increasing for $k \in (\max \{0, \cos \alpha\}, 1)$
6. if condition 2 fails, for $k \in (0, 1)$, increasing if $\delta < \frac{1}{1-r_i+r_{-i}}$, constant if $\delta = \frac{1}{1-r_i+r_{-i}}$, decreasing if $\delta > \frac{1}{1-r_i+r_{-i}}$

**Proof.** The continuity in part 1 follows from the continuity of $f$, $d(k b_i, b_i)$ and $d(k b_{-i}, b_i)$. 

36
The differentiability follows from the derivatives below, which can be easily checked:

\[
\begin{align*}
    d(kb_i, b_i) &= b|1 - k| & d(kb_{-i}, b_i) &= b\sqrt{(k - \cos \alpha)^2 + \sin^2 \alpha} \\
    d'(kb_i, b_i) &= \begin{cases} < 0 & \text{if } k \in [0, 1) \\ \not\in & \text{if } k = 1 \\ > 0 & \text{if } k \in (1, \infty) \end{cases} & d'(kb_{-i}, b_i) &= \begin{cases} < 0 & \text{if } k < \cos \alpha \\ = 0 & \text{if } k = \cos \alpha \wedge \alpha \neq 0 \\ \not\in & \text{if } k = \cos \alpha \wedge \alpha = 0 \\ > 0 & \text{if } k > \cos \alpha \end{cases} \\
    d''(kb_i, b_i) &= \begin{cases} = 0 & \text{if } k \neq 1 \\ \not\in & \text{if } k = 1 \end{cases} & d''(kb_{-i}, b_i) &= \begin{cases} 0 & \text{if } \alpha = \pi \vee (k \neq 1 \wedge \alpha = 0) \\ \not\in & \text{if } k = 1 \wedge \alpha = 0 \end{cases}
\end{align*}
\]

For parts 2 and 3, we have:

\[
\begin{align*}
    \tilde{U}'_i(k) &= f'(d(kb_i, b_i))d'(kb_i, b_i)(1 - \delta r_{-i}) + \\
    &\quad f'(d(kb_{-i}, b_i))d'(kb_{-i}, b_i)\delta(1 - r_i) \\
\end{align*}
\]

and direct verification shows that \(\tilde{U}'_i(k) > 0\) for \(k < \cos \alpha\), \(\lim_{k \to \cos \alpha} -\tilde{U}'_i(k) \geq 0\), \(\tilde{U}'_i(k) < 0\) for \(k > 1\) and \(\lim_{k \to 1\wedge} \tilde{U}'_i(k) \leq 0\).

For the strict concavity for \(k \in (\max \{0, \cos \alpha\}, 1)\) in part 4, we have

\[
\begin{align*}
    \tilde{U}''_i(k) &= f''(d(kb_i, b_i))[d'(kb_i, b_i)]^2 (1 - \delta r_{-i}) + \\
    &\quad f'(d(kb_i, b_i))d'(kb_i, b_i)(1 - \delta r_{-i}) + \\
    &\quad f''(d(kb_{-i}, b_i))[d'(kb_{-i}, b_i)]^2 \delta(1 - r_i) + \\
    &\quad f'(d(kb_{-i}, b_i))d'(kb_{-i}, b_i)\delta(1 - r_i) \\
\end{align*}
\]

\(\tilde{U}''_i(k) \leq 0\) is a consequence of the fact that all the summands of the expression are non-negative. To see that \(\tilde{U}''_i(k) < 0\) under conditions 1 and 2, note that the last summand is either zero, which implies that the next to last summand is negative, or negative.

For part 5, when condition 1 fails, \(\tilde{U}_i(k) = f(d(kb_i, b_i)) \cdot c\) where \(c > 0\) so that \(\tilde{U}_i(k)\) is increasing on \((\max \{0, \cos \alpha\}, 1)\). Finally, when condition 2 in part 6 fails, we have \(f'' = 0\) and \(\alpha = \pi\) so that \(\tilde{U}_i'(k) = 1 - \delta r_{-i} - \delta(1 - r_i)\). \(\square\)
Proposition 4 (SSMPE with General Utility). Assume, without loss of generality, that $r_1 \geq r_2$. Then:

1. if condition 1 fails, $\hat{k}_1 = \hat{k}_2 = 1$ characterize the unique SSMPE;

2. if condition 2 fails, then if $\delta < \frac{1}{1-r_2+r_1}$, $\hat{k}_1 = \hat{k}_2 = 1$ characterize the unique SSMPE, if $\delta > \frac{1}{1-r_2+r_1}$, $\hat{k}_1 = 1$ and $k_2 = 0$ characterize the unique SSMPE, and if $\delta = \frac{1}{1-r_2+r_1}$, $\hat{k}_1 = 1$ and any $\hat{k}_2 \in [0,1]$ characterize a SSMPE;

3. if conditions 1 and 2 hold, there exists a unique $\kappa_i$ for $i \in \{1, 2\}$ given by $\kappa_i = \arg\max_{k \geq 0} \tilde{U}_i(k) \in [\max\{0, \cos \alpha\}, 1]$ and either $\kappa_1 > \kappa_2$, in which case $\hat{k}_1 = 1$ and $\hat{k}_2 = \kappa_2$ characterize the unique SSMPE, or $\kappa_1 = \kappa_2$, in which case there exist exactly two SSMPE characterized by $\hat{k}_1 = 1$, $\hat{k}_2 = \kappa_2$ and $\hat{k}_1 = \kappa_1$, $\hat{k}_2 = 1$.

Proof. Clearly, if $\delta = 0$, or $r_2 = 1$ (which implies $r_1 = 1$), or $\alpha = 0$, then none of the players has any incentive to moderate. By Lemma A2 there exists a unique maximizer of $\tilde{U}_i(k)$ for $i \in \{1, 2\}$, $k = 1$, and using arguments similar to the proof of Proposition 1, there exists a unique SSMPE characterized in Proposition 4.

When condition 2 fails, then by Lemma A2, depending on the relationship of $\delta$ and $\frac{1}{1-r_2+r_1}$, the maximizer of $\tilde{U}_2(k)$ is either $k = 1$, $k = 0$ or any $k \in [0,1]$. To see that $\tilde{U}_1(k)$ is always maximized at $k = 1$, note that $r_1 \geq r_2$ implies $\frac{1}{1-r_1+r_2} \geq 1$. Using again similar argument as in the proof of Proposition 1, if $\delta \neq \frac{1}{1-r_2+r_1}$, there exists a unique SSMPE with $\hat{k}_1 = 1$ and either $\hat{k}_2 = 0$ or $\hat{k}_2 = 1$. If $\delta = \frac{1}{1-r_2+r_1}$, then any $\hat{k}_1 = 1$ and $\hat{k}_2 \in [0,1]$ characterize a SSMPE.

When both conditions 1 and 2 hold, the uniqueness of $\kappa_i = \arg\max_{k \geq 0} \tilde{U}_i(k)$ and $\kappa_i \in [\max\{0, \cos \alpha\}, 1]$ follow from Lemma A2. Recalling again the proof of Proposition 1, uniqueness/multiplicity of a SSMPE and its characterization follows. Finally, we claim that $\kappa_1 \geq \kappa_2$. If $\kappa_2 = 0$, the claim is trivial. If $\kappa_2 \in (\max\{0, \cos \alpha\}, 1)$, the claim follows from Proposition 5 that we prove below. If $\kappa_2 = 1$, then $\lim_{k \to 1^-} \tilde{U}'_2(k) \geq 0$ and we have

$$0 \leq \lim_{k \to 1^-} \tilde{U}'_2(k)$$

$$= -f'(0)(1 - \delta r_1) + f'(d(b_2, b_1)) d'(k b_2, b_1) \big|_{k=1} \delta (1 - r_2)$$

$$\leq -f'(0)(1 - \delta r_2) + f'(d(b_1, b_2)) d'(k b_1, b_2) \big|_{k=1} \delta (1 - r_1)$$

$$= \lim_{k \to 1^-} \tilde{U}'_1(k)$$

(A15)
where the inequality follows from $-f'(0) \geq 0$, $r_1 \geq r_2$, $d(b_1, b_2) = d(b_2, b_1)$ and $d'(kb_1, b_2)|_{k=1} = d'(kb_2, b_1)|_{k=1}$. From the strict concavity of $\tilde{U}_1(k)$ on $(\max \{0, \cos \alpha\}, 1)$ and $\lim_{k \to 1-} \tilde{U}_1'(k) \geq 0$ it follows that $\kappa_1 = 1$.

**Proposition 5** (SSMPE Comparative Static with General Utility).

Assume conditions 1 and 2 hold, $r_1 \geq r_2$ and $\kappa_2 \in (\max \{0, \cos \alpha\}, 1)$. Then for $\kappa_2$ in Proposition 4 (assuming $\alpha \neq \pi$ for the first relation),

$$
\begin{align*}
\frac{\partial \kappa_2}{\partial \alpha} &< 0 & \frac{\partial \kappa_2}{\partial \delta} &< 0 \\
\frac{\partial \kappa_2}{\partial r_1} &< 0 & \frac{\partial \kappa_2}{\partial r_2} &> 0.
\end{align*}
$$

(A16)

**Proof.** Under the conditions of the Proposition, $\kappa_2$ is implicitly defined by $\tilde{U}_2'(\kappa_2) = 0$. From the implicit function theorem $\frac{\partial \kappa_2}{\partial x} = \frac{\partial \tilde{U}_2'(\kappa_2)}{\partial x}$, $-\tilde{U}_2'(\kappa_2)$. The denominator of this expression is positive by strict concavity of $\tilde{U}_2(k)$. The numerator of this expression for $x \in \{\alpha, \delta, r_1, r_2\}$ is

$$
\begin{align*}
\frac{\partial}{\partial \alpha} \tilde{U}_2'(\kappa_2) &= f''(d(\kappa_2b_1, b_2))d'(\kappa_2b_1, b_2)\delta(1-r_2)\frac{\partial d(\kappa_2b_1, b_2)}{\partial \alpha} \\
&+ f'(d(\kappa_2b_1, b_2))\frac{\partial d'(\kappa_2b_1, b_2)}{\partial \alpha}\delta(1-r_2) \\
\frac{\partial}{\partial \delta} \tilde{U}_2'(\kappa_2) &= f'(d(\kappa_2b_2, b_2))d'(\kappa_2b_2, b_2)(-r_1) \\
&+ f'(d(\kappa_2b_1, b_2))d'(\kappa_2b_1, b_2)(1-r_2) \\
\frac{\partial}{\partial r_1} \tilde{U}_2'(\kappa_2) &= f'(d(\kappa_2b_2, b_2))d'(\kappa_2b_2, b_2)(-\delta) \\
\frac{\partial}{\partial r_2} \tilde{U}_2'(\kappa_2) &= f'(d(\kappa_2b_1, b_2))d'(\kappa_2b_1, b_2)(-\delta)
\end{align*}
$$

(A17)

where the first three expressions are negative and the last one positive. □