Federal Reserve Bank of Philadelphia

Working Papers
Research Department

WP 19-07
January 2019
https://doi.org/10.21799/frbp.wp.2019.07

Incumbency Disadvantage of Political Parties: The Role of Policy Inertia and Prospective Voting

Satyajit Chatterjee
Federal Reserve Bank of Philadelphia Research Department

Burcu Eyigungor
Federal Reserve Bank of Philadelphia Research Department

ISSN: 1962-5361

Incumbency Disadvantage of Political Parties: The Role of Policy Inertia and Prospective Voting

Satyajit Chatterjee and Burcu Eyigungor

Federal Reserve Bank of Philadelphia

January 2019

Abstract

We document that postwar U.S. elections show a strong pattern of "incumbency disadvantage": If a party has held the presidency of the country or the governorship of a state for some time, that party tends to lose popularity in the subsequent election. To explain this fact, we employ Alesina and Tabellini's (1990) model of partisan politics, extended to have elections with prospective voting. We show that inertia in policies, combined with sufficient uncertainty in election outcomes, implies incumbency disadvantage. We find that inertia can cause parties to target policies that are more extreme than the policies they would support in the absence of inertia and that such extremism can be welfare reducing.

Keywords: rational partisan model, incumbency disadvantage, policy inertia, prospective voting, median voter

JEL Codes: D72 H50

---

*Satyajit.Chatterjee@phil.frb.org. The authors thank Marina Azzimonti, Hülya Eraslan, Chris Phelan, Razvan Vlaicu, and Pierre Yared for very helpful comments. Comments from conference participants at the 2017 Society for Economic Dynamics Meetings, 2017 Stony Brook Summer Workshop in Political Economy, the 2017 Midwest Macro Meetings at Pittsburgh and 2018 Minnesota Workshop in Macro Theory and from seminar participants at Fordham University, Ohio State University and Rice University are gratefully acknowledged. The authors thank Matt Grossman for sharing the data for the chart from his October 4, 2018, *FiveThirtyEight* blog post.

This paper supersedes "Incumbency Disadvantage in U.S. National Politics: The Role of Policy Inertia and Prospective Voting" by Satyajit Chatterjee and Burcu Eyigungor, Federal Reserve Bank of Philadelphia Working Paper 17-43, December 2017.

1 Introduction

This paper is motivated by the following observation. Since 1954, there have been eight presidential elections in which the presidency had been held continuously by either a Democrat or a Republican for the eight preceding years or more (two or more terms). In seven of those elections, the incumbent president's party could not hold on to the presidency. This fact suggests that there is an *incumbency disadvantage* in U.S. politics: When a party has held the presidency for two or more terms, the popularity of the party with voters is strongly diminished.[1]

We make several contributions. First, we verify that the suggestion of incumbency disadvantage noted previously is in fact present in national and gubernatorial elections in the postwar era. Specifically, we show that the Democratic vote share in the House is strongly affected by how long the two parties have held the presidency going into each election: If the Democratic (Republican) Party has held the presidency for six or more years going into an election, the Democratic vote share of the House declines (increases) by 2.4 percentage points, on average.[2] We also study state gubernatorial elections, which allows us to expand the number of elections we can examine. We find that if a party has held the governor's office for six or more years, that party's candidate for governorship garners fewer votes in the subsequent election.[3]

Second, we show that incumbency disadvantage is implied by the (Markov perfect) equilibrium of Alesina and Tabellini's (1990) model of partisan politics, if their model is extended in two very natural ways. First, political turnover occurs via elections in which outcomes depend on the anticipated policy choices of the two parties as well as on transient voter preference shocks, and, second, there is policy inertia stemming from the costs of (or constraints on) changing policies quickly. A simple numerical example can illustrate that both inertia in policies and diminishing marginal utility play crucial roles in getting incumbency disadvantage in the model. Suppose that the Democratic Party's preferred combination of expenditure is \$100 on food stamps and \$0 on defense, and the Republican Party's preferred combination of expenditures is \$0 on food stamps and

[1] After their first term in office, most presidents get reelected (this fact may reflect the personal appeal of a president once the public gets to know him and is not addressed in this paper). After two terms in office, a president cannot run for a third term, so the identity of the next presidential party mostly depends on the appeal of party platforms.

[2] We focus on the House because every seat in the House is generally contested in every national election (held every two years), and the scope of the electorate to express approval or disapproval of current policies is the greatest in House elections.

[3] In a political system such as in the U.S., the president's and governor's party gets to set the policy agenda at the national and state levels, respectively. So, during national and state elections, we expect voters to vote against (or for) the members of the president's or governor's party if they disapprove (or approve) of the party's current policies.

$100 on defense. Suppose the Democratic Party has been in power a long time, so the composition of spending is at the party's ideal combination. If the Democratic Party is reelected, the composition of spending will not change. If the Republican Party wins, then, due to inertia in policies, the composition will change to $10 on defense and $90 on food stamps. When there is diminishing marginal utility, this $10 increase in defense is much more valuable to Republican voters than the $10 loss in food stamps is to Democratic voters. As a result, in the model, the preference shocks will be less determinative for Republicans, and they are more likely to vote along partisan lines and win the election. The same logic applies in reverse after a long Republican incumbency.

Third, we show that the policy inertia needed to account for incumbency disadvantage has important positive and normative implications. On the positive side, each party's long-run ideal policy may become more extreme than the incumbent party member's static ideal policy (i.e., the policy that maximizes the party member's period utility). This is significant because when such an extreme policy is followed, a less extreme policy would improve *all* voters' period utility. Since inertia leads to incumbency disadvantage, parties target more extreme policies so they can enjoy policies closer to their ideal even when the opposition is calling the shots. On the normative side, this extremism may entail a welfare loss for all voters.

Fourth, we show that policy inertia and prospective voting enlarge the set of environments in which Downsian competition between two parties that care about winning the election leads to centrist policies being adopted. As Alesina (1988) first pointed out, if the two parties have their own agendas and neither party can commit to policies before the election, then in a Markov equilibrium, the winning party will implement its preferred, not median, policy *regardless* of its desire to hold office (i.e., regardless of the intensity of office motivation).[4] In our model as well, without inertia and preelection commitment, the Markov equilibrium will have parties choosing their preferred policies regardless of their desire to hold office (i.e., there will be no movement toward centrist policies stemming from a desire to win and retain elected office). But policy inertia and prospective voting change this result, and office motivation begins to matter for policy choice: Choosing policies closer to voters' ideal policies means a higher chance of losing the next election and, thus, a higher chance of losing the utility from being in office. The additional private cost of

[4]If parties can commit to a platform before the election, the Downsian outcome becomes possible even with candidates who compete because they care about policies (Wittman (1983), Calvert (1985)). The Downsian outcome can also be resurrected if purely ideologically motivated parties play a repeated game with trigger strategies, as shown in Alesina (1988).

extremism pulls equilibrium policies closer to the center. Thus, the Downsian logic remains active even when parties are ideologically motivated and there is no preelection commitment.

Finally, our paper makes a methodological contribution by augmenting the toolkit of researchers using quantitative theory to study partisan politics with endogenous reelection probabilities. Such games can be hard to compute because of the possible lack of continuity of Markovian decision rules (Chatterjee and Eyigungor (2016), Cao and Werning (2018)) and, consequently, the possible nonexistence of a stationary pure strategy equilibrium. By incorporating a continuously distributed i.i.d. preference shock, we ensure the existence of a Markov perfect equilibrium in pure strategies. The existence result puts our model — and other models that share key features of our environment — on a secure computational footing.[5]

Our paper is related to several literatures. On the empirical side, our finding of incumbency disadvantage echoes previous findings in the politics literature. In an early study, Stokes and Iverson (1962) observed that, over the 24 presidential elections between 1868 and 1960, neither the Republican nor the Democratic Party succeeded in winning more than 15 percent beyond an equal share of presidential or congressional votes. They interpret their findings as evidence of "restoring forces" that work to elevate the popularity of the party that has been less popular in the past. More recently, Bartels and Zaller (2001) and Fair (2009) study a large set of empirical presidential vote models and identify an"incumbent fatigue" effect (Bartels and Zaller) and "duration" effect (Fair), wherein the percent of the two-party vote for the party of the incumbent president is negatively affected by how long that party has held the presidency. Our findings for gubernatorial elections are similar to theirs. Relatedly, Erikson (1988) and Alesina and Rosenthal (1995) noted that the president's party loses seats in midterm elections (the so-called midterm cycle). We show that there is a long-term incumbency disadvantage not related to midterm cycles.

In the area of macro political economy, we contribute to the growing literature on quantitative-theoretic partisan political economy models featuring endogenous reelection probabilities. In terms of model structure and quantitative focus, the closest is Azzimonti (2011), which also applies prob-

[5]The potential lack of continuity of decision rules and, therefore, of the mapping whose fixed points are pure strategy Markov equilibria also plagues models of industry dynamics (Doraszelski and Satterthwaite (2010), models of legislative bargaining (Duggan and Kalandrakis (2012)) and models of sovereign debt and default (Chatterjee and Eyigungor (2012)). In all instances, the existence problem is solved by introducing i.i.d. shocks in the right places. We note that the existence proof given in Duggan and Kalandrakis (2012) cannot be used to claim existence for the present model since a key element of the present model — endogenous reelection probability — is missing in theirs.

abilistic voting (Lindbeck and Weibull (1987)) to a dynamic setup.[6] However, while the reelection probability is endogenous in Azzimonti's model, her assumption regarding preferences (and the symmetry of the model) implied that the probability of reelection is *state independent*. Thus, by construction, her model is silent on how reelection probabilities evolve over time. Prospective voting also features importantly in Alesina and Rosenthal's (1995) explanation of the midterm cycle. Their explanation centered around nonpolarized voters attempting to get closer to median policies by counterbalancing a partisan president's power. In our approach, citizens are as polarized as parties, but for random reasons, they do not always vote along party lines. Coupled with the policy drift resulting from inertia, we predict an electoral disadvantage that grows with incumbency. Prospective voting and endogenous reelection probability also feature in Rogoff's (1990) model of the political budget cycle. The friction in his model is asymmetric information about the competency of the incumbent leader and the cycle arises from the competent type choosing policies that separate him or her from the incompetent type. There are cycles in taxes and expenditure policies but not necessarily incumbency disadvantage. Prospective voting and endogenous reelection probabilities have recently featured in quantitative models of sovereign debt and default. Scholl (2017) uses a version of Persson and Svensson's (1989) partisan politics model to quantitatively explore the implications of endogenous reelection probabilities on sovereign borrowing and default behavior, and Chatterjee and Eyigungor (2017) study the role of prospective voting in accounting for observed fluctuations in the risk of sovereign default in emerging economies. But incumbency disadvantage is not a necessary feature of these models.

Endogenous political turnover is a feature of Ales, Maziero, and Yared's (2014) model of political cycles. These authors approach political turnover as the outcome of a principal/agent contracting problem (citizens as principal and the government as agent). Because the agent (government) has private information about the state of the budget, the optimal contract has the feature that, after a sequence of bad outcomes, the contract is terminated and a new contract is entered into with a new agent. Replacement of the government is, thus, endogenous, and the logic is closely tied to that in models of political control and retrospective voting (Barro (1973) and Ferejohn (1986)). But there is no incumbency disadvantage. Cycles in expenditures tied to exogenous shifts in the party or coalitions in power occur in Dixit, Grossman, and Gul (2000), Acemoglu, Golosov, and Tsyvinski

[6] Other recent examples of quantitative political economy models include Mateos-Planas (2012) and Song, Storesletten, and Zilibotti (2012).

(2011), Aguiar and Amador (2011), Battaglini and Coate (2008), and Azzimonti, Battaglini, and Coate (2016).

Our results on the positive and normative implications of the costs of adjustments bear a resemblance to equilibrium outcomes in legislative bargaining models with an endogenous status quo (Bowen, Chen, and Eraslan (2014), Piguillem and Riboni (2015) and Dziuda and Loeper (2016)). For instance, Bowen, Chen, and Eraslan (2014) show that mandated spending improves the bargaining position of politicians out of power and leads to equilibrium outcomes closer to the first best.[7] A different strand of the literature (Aghion and Bolton (1990); Milesi-Ferretti and Spolare (1994); Besley and Coate (1998); Hassler, Mora, Storesletten, and Zilibotti (2003)) finds, as we do, the possibility of inferior policies being adopted by incumbent governments (or superior policies ignored) for strategic electoral reasons.

The paper is organized as follows. In Section 2, we present our empirical findings regarding incumbency disadvantage of political parties. In Section 3, we develop the model outlined previously. In Section 4, we analyze a static version of the model to provide intuition explaining why policy inertia leads to incumbency disadvantage. In Section 5, we explore the full dynamic model computationally — the goal of this section is to show that analogs of the empirical findings reported in Section 2 can arise in the full equilibrium of the model and to understand the key factors that lead to this result. In Sections 6 and 7, we examine the positive and normative implications of policy inertia and incumbency disadvantage, and Section 8 concludes. Appendix A provides some additional regression results to buttress the empirical findings reported in Section 2. Appendix B gives the proof of the existence of a pure strategy Markov perfect equilibrium, and Appendix C provides an algorithm to compute the (potentially discontinuous) Markovian decision rules of the two parties.

2 Incumbency Disadvantage in U.S. Elections: 1954–2016

In this section, we show that the incumbency disadvantage effect mentioned in the opening paragraph is a feature of postwar U.S. election outcomes. As noted in our introduction, in empirical models designed to predict presidential election outcomes, the duration of the presidential party's

[7] The dynamic link created by mandated spending is analogous to the dynamic link created by the costs of changing inherited policies in our model. However, an important reason why this dynamic link matters in our model is that it endogenizes reelection probabilities. In contrast, in Bowen, Chen, and Eraslan's (2014) environment (and in environments of legislative bargaining models in general), the probability of a legislator being chosen to propose a policy is exogenously given and does not depend on the policies the legislator is expected to propose.

incumbency negatively affects the party's reelection probability. In the first subsection, we evaluate party performance using U.S. House election outcomes, which doubles the number of elections we can consider and has the advantage that the election outcome is less affected by the personal appeal of the two (presidential) candidates.[8] In the second subsection, we extend the analysis to state gubernatorial elections, which allows us to further expand the number of elections we can examine. Our results here parallel the findings for presidential elections: If a party holds the governor's office for a long time, that party's candidate for governor garners fewer votes.

2.1 Congressional Elections

Letting DV denote the percentage two-party share of votes garnered by House Democrats in each national election, our main empirical specification is:[9]

$$DV_t = \beta_0 + \beta_1 MIDTERM_t + \beta_2 SIX_t^+ + \beta_3 (TWO_t^+ \cdot RDPIGR_t).$$

Here, SIX_t^+ is a trichotomous variable that takes a value of +1 if, at the time of the election, the presidency has been held by a Democrat for six or more years, takes a value of -1 if the presidency has been held by a Republican for six or more years and takes the value 0 otherwise. If there is an incumbency disadvantage, we expect β_1 to be negative. A negative coefficient implies that after six years of a Democratic presidency, the Democratic vote share falls, and after six years of a Republican presidency, the Democratic vote share rises.

$RDPIGR_t \cdot TWO_t^+$ is an interaction variable, where TWO_t^+ is a binary variable that takes a value of 1 if the presidency was held by a Democrat in the preceding two or more years, and it takes the value -1 if the presidency was held by a Republican in the preceding two or more years. And $RDPIGR_t$ is the deviation from the sample mean of the growth rate of real disposable per-capita income from the third quarter of the previous year to the third quarter of the election year. This interaction term takes into account that above-average economic performance in the preceding year

[8]The drawback of looking only at presidential vote shares is that, once we take into account incumbency advantage at the individual politician level, it does not leave many elections (in the modern era) in which we can evaluate the popularity of the presidential party: In our sample, there are only six presidential elections in which the incumbent president is not running again.

[9]The data on House vote share and House vote seats are compiled from the official website of the U.S. House of Representatives at $https://history.house.gov/Institution/Party-Divisions/Party-Divisions/$. The data for presidential party incumbency are from $https://en.wikipedia.org/wiki/President\_of\_the\_United\_States$. The data for real personal disposable income are from the U.S. Bureau of Economic Analysis, Real Disposable Personal Income: Per Capita [A229RX0], retrieved from FRED, Federal Reserve Bank of St. Louis; $https://fred.stlouisfed.org/series/A229RX0$.

may be attributed to the success of policies of the presidential party, and so, the presidential party gains more votes. If so, we expect β_3 to be positive.

$MIDTERM_t \cdot TWO_t^+$ is also an interaction variable, where $MIDTERM_t$ is a dummy variable that is 1 for midterm elections and 0 otherwise. The variable takes into account the tendency of voters to balance the power of a newly elected or newly reelected partisan president by electing representatives from the nonpresidential party in greater numbers (Erikson (1988), Alesina and Rosenthal (1995)), and so β_2 is predicted to be negative.

Table 1 presents our estimation results.[10] The first column reports the results of the main regression. The β_2 coefficient is estimated to be negative and statistically significant at the 1 percent level. Evidently, a long presidential incumbency is costly for the party: On average, a party's vote share declines by 2.37 percentage points in elections in which the party held the presidency for six or more years. The β_3 coefficient is estimated to be positive and statistically significant at the 5 percent level, which accords with the common finding (see, for instance, Bartels and Zaller (2001) and Fair (2009)) that good economic performance boosts the popularity of the president's party. The β_4 midterm coefficient has the predicted sign, but the coefficient is not estimated to be statistically significant.

The second column checks for robustness with respect to the measure of economic performance by using per-capita growth in real GDP instead of per-capita growth in real disposable income. The magnitudes of the estimated coefficients remain very similar.

While our main regression specification follows the literature in ignoring the bounded nature of the vote share variable, we confirm in the third column that the relationships hold when DV is replaced with $\ln[DV/(1-DV)]$, which is an unbounded variable.

Finally, in Appendix A, we report the estimation results of our main regression using the Democratic share of House seats as the dependent variable. The coefficient on the SIX^+ variable remains significant at the 1 percent level.

How robust is the finding of incumbency disadvantage to the length of incumbency? Table 2 reports estimates of β_2 as the length of incumbency of the presidential party is varied from two or more years to eight or more years. In all cases, the β_2 coefficient is estimated to be negative,

[10]In this estimation, we are following the common practice (see, for instance, Lewis-Beck and Rice (1984) and Campbell (1997) among others) of ignoring that vote shares are bounded variables.

Table 1:
Presidential Incumbency of a Party and Democratic Share of House Votes

DEP. VAR.	DV	DV	$\ln(DV/(1-DV))$
CONSTANT	51.75***	51.68***	0.07**
	(0.66)	(0.65)	(0.03)
SIX^+	−2.37***	−2.49***	−0.10***
	(0.63)	(0.63)	(0.03)
$PDPIGR * TWO^+$	0.40**		0.02**
	(0.20)		(0.01)
$PGDPGR * TWO^+$		0.31*	
		(0.15)	
$MIDTERM * TWO^+$	−1.01	−0.93	−0.04
	(0.89)	(0.89)	(0.04)
SD(DEP VAR)	3.22	3.22	0.13
NO. OF OBS.	32	32	32
R^2	0.51	0.49	0.51
ADJ. R^2	0.45	0.44	0.45

Table 2:
Presidential Incumbency of a Party and Democratic Share of House Votes
Other Specifications

DEP. VAR.		*DV*		
CONSTANT	51.90***	51.83***	51.750***	51.79***
	(0.50)	(0.67)	(0.66)	(0.65)
TWO^+	−1.08			
	(0.69)			
$FOUR^+$		−1.30***		
		(0.42)		
SIX^+			−2.37***	
			(0.63)	
$EIGHT^+$				−2.21***
				(0.50)
$PDPIGR * TWO^+$	0.41	0.44*	0.40**	0.36
	(0.25)	(0.22)	(0.205)	(0.23)
$MIDTERM * TWO^+$	−1.10	−1.53**	−1.01	−2.07***
	(0.97)	(0.73)	(0.89)	(0.71)
SD DEP. VAR.	3.22	3.22	3.22	3.22
NO. OF OBS.	32	32	32	32
R^2	0.36	0.40	0.51	0.43
ADJ. R^2	0.29	0.34	0.45	0.37

and it is strongly statistically significant (at the 1 percent level) when the length of incumbency is four or more years, six or more years (the main regression), and eight or more years. Furthermore, the estimated β_2 coefficient is similar in magnitude for incumbency length of six or more years and eight or more years, and the adjusted R^2 is highest for the main regression. The coefficients on the two control variables (β_3 and β_4) are always estimated to have the predicted signs, but their statistical significance varies.

The message we take from Tables 1 and 2 is that there is robust evidence that a party that has held the presidency for some length of time suffers a loss in its popularity.

2.2 Gubernatorial Elections

In this section, we investigate whether the incumbency disadvantage effect is also operative for elections of state governors. The idea here parallels the extant investigations done for U.S. presidential elections. Within a state, the governor holds agenda-setting powers that can be used to further the policy program of his or her party. Our goal is to determine whether the popularity of gubernatorial candidates is lessened if the candidates' party has held the governorship for a long time.

For this investigation, we examined the two-party share of the gubernatorial vote in all elections for the 50 states between 1954 and 2016 subject to data availability.[11] We limit consideration to those elections in which the two parties are competitive. We do this by excluding the initial years in a state until both parties have won at least one election in the preceding 20 years. From that year on, we include all gubernatorial elections in that state.[12] In addition, we ignore elections in which the top two parties in the election outcome are not Democrats and Republicans.

For our main regression, we restrict attention to gubernatorial elections in which the incumbent candidate faces a binding term limit and therefore cannot run for reelection.[13] This restriction serves to ensure that a finding of incumbency disadvantage is more likely to reflect the (un)popularity

[11]The data on governor races are compiled from *https : //www.ourcampaigns.com/*. State-level real per-capita income comes from the U.S. Bureau of Economic Analysis.

[12]The starting year in the sample for the affected states that were historically heavily Democratic are: Alabama (1990), Arkansas (1968), Florida (1970), Georgia (2006), Mississippi (1995), North Carolina (1976), Oklahoma (1966), Tennessee (1974), Texas (1982), and Virginia (1973). The starting year in the sample for the affected states that were historically heavily Republican are: Minnesota (1956), New Hampshire (1964), and Vermont (1964).

[13]Whether the incumbent governor faced a term limit was determined on a case-by-case basis from various websites.

of the incumbent governor's *party* rather than of the governor him or herself.[14] However, this restriction is costly in that we can use only one-fifth of our observations. So, we also examine the outcome of all gubernatorial elections in our sample, with a control put in for whether an incumbent governor is running for reelection.

Letting GDV denote the deviation from the state sample mean of the vote share of Democratic gubernatorial candidates, our main specification is the panel regression

$$GDV_t = \beta_1 SIX_t^+ + \beta_2 PPIGR_t \cdot TWO_t^+ + YEAR \text{ fixed effects.}$$

Here, SIX_t^+ has the same interpretation as in the congressional elections: It takes the value +1 if Democrats have held the governor's office for six or more years, −1 if the Republicans have held the office for six or more years, and 0 otherwise. $PPIGR_t \cdot TWO_t^+$ is an interaction variable, where $PPIGR_t$ is the deviation from the state sample mean of the growth in state per-capita disposable income in the four quarters preceding the election quarter and TWO_t^+ is a binary variable that takes the value +1 if the Democratic Party held the governor's office in the two years preceding the election and −1 if the Republican Party did so. We include year fixed effects to capture nationwide swings in favor of or against the two parties. Since the dependent variable is demeaned, state fixed effects decrease adjusted R^2 and are not included. As in the congressional regressions, β_1 is predicted to be negative and β_2 is predicted to be positive.

The first two columns of Table 3 report the estimation results for our main specification. The first column is the restricted sample in which the incumbent governor cannot run for reelection again. We find a highly significant negative effect of incumbency, confirming that incumbency disadvantage is evident in gubernatorial elections as well. The coefficient on state-level economic performance is positive but only mildly statistically significant. The results for the unrestricted sample appear in the second column. For this regression, we include a dummy variable INC_t that turns on if the incumbent governor is running for reelection. Once again, there is strong statistical evidence of incumbency disadvantage of the party, with the magnitude of the effect being somewhat smaller than in the main regression.[15] The regression also shows that an incumbent governor enjoys

[14]Incumbency advantage of an incumbent politician running for reelection is widely established for the U.S. (see, for instance, Erikson (1971), Gelman and King (1990), Ansolabehere and Snyder Jr. (2002), Mayhew (2008), and Jacobson (2015)).

[15]Term limits for governors are more common in the later part of our sample period and the two parties are also more competitive in all states in the later part of the sample period. Thus, we would expect the incumbency

a strong (in magnitude as well as statistical significance) incumbency *advantage*, consistent with a large existing literature on incumbency advantage for individual politicians.

Table 3:
Gubernatorial Incumbency of a Party and Its Share of Gubernatorial Votes

DEP. VAR.	*GDV*		*GDVN*	
SIX^{+}	−2.24***	−1.49***	−5.04***	−5.12***
	(0.60)	(0.44)	(0.92)	(0.47)
$PPIGR \cdot TWO^{+}$	0.33	−0.04	.49	0.02
	(0.23)	(0.08)	(0.34)	(0.10)
INC		6.69***		7.68***
		(0.40)		(0.47)
YEAR FE	YES	YES	YES	YES
STATE FE	NO	NO	YES	YES
SD DEP. VAR.	6.21	9.23	9.10	10.40
NO. OF OBS	157	782	153	770
R^2	0.57	0.40	0.80	0.42
$ADJ.R^2$	0.31	0.34	0.52	0.32

The final two columns of Table 3 consider a detrended version of our dependent variable. Although the two parties are competitive in national elections in the modern era, there is well-known geographic variation in the popularity of the two parties that is quite persistent through time. We would expect geographic persistence of party popularity to dampen the operation of the incumbency disadvantage effect at the state level. To address this, we follow Tufte (1978) and examine the impact of our explanatory variables on the *deviations* of vote shares from "normal" vote shares. Tufte defines the "normal" vote share of a party in any given election as the average vote share of the party in the previous 16 years (i.e., its average in eight previous elections if elections happen every two years and its average over four previous elections if elections happen

disadvantage effect to be more active and, therefore, the magnitude of the β_1 coefficient to be higher in the restricted sample.

every four years). The dependent variable, denoted $GDVN$, is then the deviation of the two-party vote shares of Democratic gubernatorial candidates from the normal two-party vote shares of Democratic gubernatorial candidates in the state. Since the deviations from normal vote shares need not have zero mean, these regressions also include STATE fixed effects.

The results show that the estimated incumbency disadvantage effect remains statistically highly significant and is now much larger in magnitude. Indeed, all coefficients that are estimated to be statistically significant are larger in magnitude and all coefficients have the predicted signs. These results are consistent with the expectation that deviations of vote shares from the "normal" vote share should be more responsive to variations in explanatory factors.

To summarize, the results in Table 3 show that the incumbency disadvantage effect is also present in governor elections in the post-1954 era. In Appendix A, we buttress this conclusion by showing that these findings are robust with respect to a shorter definition of duration of incumbency.

2.3 Do Policy Choices Matter in Elections?

As noted in the introduction, our explanation of incumbency disadvantage works through policy choices: The ruling party implements policies that conform to its ideal policies, which then reduces the party's support among members of the other (opposition) party. For this mechanism to be active, it is necessary that the ideal policies of the two parties be different and that voters care about these differences sufficiently for the choice of policies to matter for election outcomes.

Regarding differences in the ideal policies of the two parties, Poole and Rosenthal (1985) showed that congressional roll call votes of individual legislators can be explained in terms of each member having some ideal point on a single liberal-conservative line, with Democrats occupying the liberal end of the line and Republicans, the conservative end. Many studies that followed in the wake of their path-breaking work have confirmed this finding. Since liberalism and conservatism lead to different policy choices, the first requirement of different policy agendas across the two parties seems to be satisfied.

Regarding the second requirement — that voters care about policy choices of the two parties sufficiently for these choices to matter for election outcomes — there is also compelling, if less extensive, evidence. The difficulty here is that Congress passes hundreds of laws each year, and it is challenging to determine how all this congressional activity affects election outcomes. However, in another oft-cited work, Erikson, Mackuen, and Stimson (2002) compiled a list of significant legislation passed by Congress each year and coded each such legislation as conservative or liberal or neither in character. Matt Grossman[16] compiled an updated version of this list that covers 1954–2014. The correlation between the net number of liberal laws passed in the preceding two years and the change in the House Democratic vote share from the previous election is -0.47.[17] At a broad level, this finding is consistent with our claim that, as policies move toward the ideal policies of a party, the party's popularity in elections is diminished.

More direct evidence of the loss of party popularity from enacting policies that are strongly opposed by the other party is given in recent studies that sought to understand why Democrats lost a much higher-than-predicted number of House seats in the 2010 midterm elections. The 111th Congress (President Obama's first term) saw the Democratic Party enjoying majorities in both chambers. The party pursued an ambitious policy agenda, but that agenda had no support from the Republicans. Brady, Fiorina, and Wilkins (2011, Table 2, p. 248) examined the relationship between the vote share of Democratic candidates running for reelection in 2010 and the candidates' support for the Affordable Care Act (ACA; which did not receive a single Republican vote). They find that a candidate's support for ACA reduced his or her vote share significantly, and the reduction was more severe in districts in which fewer people had voted for Obama. Kroger and Lebo (2012, Figure 5, p. 942), using a more sophisticated empirical approach, report similar results for the ACA. These findings support the claim that policy choices have electoral consequences and, again, at a broad level, our claim that as policies move toward the ideal policies of a party, the party's popularity in elections is diminished.

3 Model

3.1 Environment

We build on Alesina and Tabellini's (1990) influential model of two parties with different policy preferences circulating in power. The main differences are that election outcomes are determined endogenously via probabilistic voting and policies change inertially.

[16] "Voters Like a Political Party Until It Passes Laws," *FiveThirtyEight (blog)*, October 4, 2018, https://fivethirtyeight.com/features/voters-like-a-political-party-until-it-passes-laws/.

[17] Net liberal laws passed is the total number of liberal laws passed by a Congress minus the total number of conservative laws passed by the same Congress.

Time is discrete and denoted by $t = 0, 1, 2, \ldots$. The economy is populated by a continuum of infinitely lived individuals who derive utility from two types of public goods and discount future utility flows at the rate $\beta \in (0, 1)$. The total resources available each period to be spent on the public goods are constant $\tau > 0$. If $0 \le g_t \le \tau$ is spent on the first good in period t, then $0 \le \tau - g_t \le 1$ is spent on the second good (i.e., the transformation between the two goods is one-for-one). Although we refer to g and $\tau - g$ as public goods, g could also be interpreted as the ideological stance of policies, in which case 0 and τ would represent opposite ends of the ideological spectrum.

People's preferences toward the two public goods vary. Let $\alpha \in [0, 1)$. Then, for one-half of the population, the period utility flow from public expenditures on the two goods is

$$U(g_t, m_{1t}) + \alpha U(\tau - g_t, m_{2t})$$

while for the other half, it is

$$\alpha U(g_t, m_{1t}) + U(\tau - g_t, m_{2t}).$$

Thus, one-half of the population cares more about the first good, while the remaining one-half cares more about the second good. Here m_{1t} and m_{2t} are preference shocks that are drawn independently each period from a continuous probability distribution with support M and $U(\cdot, m_\ell) : \mathbb{R}^+ \times M \to \mathbb{R}$ is continuous, differentiable, and strictly concave for all m_ℓ, $\ell = 1, 2$. Individuals who care more about the first good are labeled D types, and those who care more about the second good are labeled R types.

The different preferences toward the two public goods motivate the political structure assumed in the model. Corresponding to the two types, there are two political parties: the D party and the R party. Each period, the two parties contest elections in which all individuals vote. The party that garners more than 50 percent of the votes wins and gets to choose the composition of public spending for that period.

A party's preferences over the two goods overlap with the type of individual it represents, but it is not identical. The period utility of the D party is

$$U(g_t, m_{1t}) + \alpha U(\tau - g_t, m_{2t}) - \psi(g_{t-1}, g_t),$$

where ψ is a cost of adjusting policies. Analogously, the period utility accruing to the R party is

$$\alpha U(g_t, m_{1t}) + U(\tau - g_t, m_{2t}) - \psi(g_{t-1}, g_t).$$

The ψ function is the way policy inertia enters the model. In reality, inertia in policies may arise from various sources. One source is that Supreme Court and federal court justices are proposed and approved by the incumbent party, which creates inertia in court decisions in favor of the policies of one party over the other. Second, once policies are put in place, they create their own vested interests that might make a reversal difficult.[18] We explore three different forms of the ψ function to show that what matters is inertia and not how it comes about. In the simplified two-period model of the next section, ψ has a simple quadratic adjustment cost form, namely, $\eta(g_{t-1} - g_t)^2$, $\eta > 0$ which gives us analytical results. In the quantitative section, we explore two more. In one, is assumed to be a step function that is 0 for $|g_{t-1} - g_t| < \Delta$ and ∞ otherwise; in this form, ψ captures the fact that there is an inherent limit to the speed with which new legislation can be implemented. In the second, the quadratic cost form is retained, but the cost is experienced by the party in power only.

The final element of the model is elections. People cast their votes in favor of the party whose policies give them the highest expected lifetime utility. Following the probabilistic voting literature, it is assumed that the elected party will make some decisions separate from the composition of public goods. An individual's preference toward these other decisions is captured by his or her *net* preference for the D party, and this net preference is the sum of an idiosyncratic component e and an aggregate component A. The components are independently drawn each period from probability distributions with CDF $F(e) : \mathbb{R} \to [0, 1]$ and $H(A) : \mathbb{R} \to [0, 1]$, respectively, with both distributions having zero mean. The idiosyncratic component is also drawn independently across individuals. Then, given e_t and A_t, a type D individual's period utility from public goods and the identity of the ruling party is

$$U(g_t, m_{1t}) + \alpha U(\tau - g_t, m_{2t}) + \mathbb{1}_{\{D \text{ party is elected in } t\}}[e_t + A_t]$$

where $\mathbb{1}_{\{\cdot\}}$ is an indicator function. A D type individual will prefer the composition of the public goods offered by the D party but may nevertheless vote for the R party if her net preference toward

[18] For instance, between 2016 and 2018, only some portions of the Affordable Care Act could be reversed.

the D party itself is sufficiently negative because of the other decisions the party will make if it comes to power. Analogously, a type R individual's period utility is

$$\alpha U(g_t, m_{1t}) + U(\tau - g_t, m_{2t}) + \mathbb{1}_{\{D \text{ party is elected}\}}[e_t + A_t],$$

and an R type may vote for the D party if her net preference for the D party is sufficiently positive.

The timeline of events in a period is as follows. At the start of the period, e shocks for all individuals and the aggregate A shock are realized. Elections are held, and each individual votes for the party that gives his or her the highest expected lifetime utility. Following the elections, the winning party makes its policy choice, taking into account the current period realization of its preference shocks. The consumption of the public goods takes place and the period closes.

3.2 Recursive Formulation

The goal of this section is to develop a recursive formulation of the model that is amenable to computation. With this goal in mind, the model will be simplified in one respect: In any period, it is assumed that only the preference shock to the preferred good of the party that wins the election is active. That is, if the D party wins the election, m_2 is set to 0 and, symmetrically, if the R party wins the election, then m_1 is set to zero.[19]

Let $g \in [0, \tau]$ denote the current period's status quo policy (i.e., the policy choice made in the previous period), and $g' \in [0, \tau]$ denote the policy choice in the current period.

Let $\Pi(g) : [0, \tau] \to [0, 1]$ denote the function that gives the probability of the D party winning the election when the status quo policy is g. The parties and the people take this function as parametrically given when making policy and voting decisions, respectively, but the function is an equilibrium object (i.e., the probabilities are required to be consistent with the equilibrium behavior of parties and voters).

[19]The computations require that there be some randomness in the choice of g' conditional on g. Having only one of the two shocks being active is enough for this purpose. If we had only m_1 or only m_2 active every period, the parties would no longer be symmetric.

We begin with the decision problem of the parties. For $k \in \{D, R\}$, let V_k denote the value of party k when it is in power and let X_k be its value when it is not in power. Then,

$$V_D(g, m_1) = \max_{g' \in [0,\tau]} U(g', m_1) + \alpha U(\tau - g') - \psi(g, g') + \beta \begin{bmatrix} \Pi(g')\mathbb{E}_{m_1'} V_D(g', m_1') + \\ [1 - \Pi(g')]\, \mathbb{E}_{m_2'} X_D(g', m_2') \end{bmatrix}. \quad (1)$$

When the D party is in power, it chooses g' taking into account the preference shock m_1 and the costs of changing policies $\psi(g', g)$. The party recognizes that its choice of g' will affect its probability of reelection next period via the function $\Pi(g')$. Let $G_D(g', m_1)$ denote a policy function that attains $V_D(g, m_1)$.

When the D party is not in power, it does not make any choices but must live with the choices made by the R party. Let $G_R(g, m_2)$ denote the policy function of party R. Then, party D's value when not in power is

$$X_D(g, m_2) = U(g') + \alpha U(\tau - g', m_2) - \psi(g, g') + \beta \begin{bmatrix} \Pi(g')\mathbb{E}_{m_1'} V_D(g', m_1') + \\ +[1 - \Pi(g')]\mathbb{E}_{m_2'} X_D(g', m_2') \end{bmatrix} \quad (2)$$
$$\text{s.t. } g' = G_R(g, m_2).$$

Symmetrically,

$$V_R(g, m_2) = \max_{g' \in [0,\tau]} \alpha U(g') + U(\tau - g', m_2) - \psi(g', g) + \beta \begin{bmatrix} \Pi(g')\, \mathbb{E}_{m_1'} X_R(g', m_1') + \\ [1 - \Pi(g')]\, \mathbb{E}_{m_2'} V_R(g', m_2') \end{bmatrix}, \quad (3)$$

and

$$X_R(g, m_1) = \alpha U(g', m_1) + U(\tau - g') - \psi(g', g) + \beta \begin{bmatrix} \Pi(g')\, \mathbb{E}_{m_2'} X_R(g', m_2') + \\ [1 - \Pi(g')]\, \mathbb{E}_{m_1'} V_R(g', m_1') \end{bmatrix} \quad (4)$$
$$\text{s.t. } g' = G_D(g, m_1).$$

Next, we turn to the value functions of people. For $k \in \{D, R\}$, let W_k be the value of type k person when her party is in power and let Z_k be her value function when her party is out of power.

Then,

$$W_D(g, m_1, e, A) = U(g', m_1) + \alpha U(\tau - g') + e + A + \beta \mathbb{E}_{(e',A',m_1',m_2')} \begin{bmatrix} \Pi(g') \, W_D(g', m_1', e', A') + \\ [1 - \Pi(g')] \, Z_D(g', m_2') \end{bmatrix} \tag{5}$$

s.t. $g' = G_D(g, m_1)$.

The value W_D depends on the status quo policy g and the D party's preference shock m_1 because these determine the policy chosen by the D party when it is in power. The value also depends on $e + A$ because of the other policies the D party will implement if it is in power. Looking forward, today's value of W_D takes into account the probability of the D party being reelected, conditional on its choice of g'.

The value to a type D individual when her party is out of power is given by

$$Z_D(g, m_2) = U(g') + \alpha U(\tau - g', m_2) + \beta \mathbb{E}_{(e',A',m_1',m_2')} \begin{bmatrix} \Pi(g') \, W_D(g', m_1', e', A') + \\ [1 - \Pi(g')] \, Z_D(g', m_2') \end{bmatrix} \tag{6}$$

s.t. $g' = G_R(g, m_2)$.

These recursions can be written more compactly. Observe that (5) implies $W_D(g, m_1, e, A)$ is additive in e and A, i.e., $W_D(g, m_1, e, A) = W_D(g, m_1, 0, 0) + e + A$. Denote $W_D(g, m_1, 0, 0)$ by $W_D(g, m_1)$, which is the value of type D when her party is in power, ignoring her current net preference for the D party. Then,

$$W_D(g, m_1, e, A) = W_D(g, m_1) + e + A. \tag{7}$$

Substituting $W_D(g, m_1) + e + A$ for $W(g, m_1, e, A)$ on the l.h.s. of (5) and substituting $W_D(g', m_1') + e' + A'$ for $W(g', m_1', e', A')$ in r.h.s. of (5) and (6) and using $\mathbb{E}_{(e',A',m_1',m_2')} e' = \mathbb{E}_{e'} e' = 0$ (the first equality follows from independence of the shocks and the second from $\mathbb{E}e = 0$) yields

$$W_D(g, m_1) = U(g', m_1) + \alpha U(\tau - g') + \beta \left[\begin{array}{c} \Pi(g')\,\mathbb{E}_{m_1'} W_D(g', m_1') + \mathbb{E}_{A'} A' | (D \text{ party win}) + \\ [1 - \Pi(g')]\,\mathbb{E}_{m_2'} Z_D(g', m_2') \end{array} \right] \tag{8}$$

$$\text{s.t. } g' = G_D(g, m_1),$$

and

$$Z_D(g, m_2) = U(g') + \alpha U(\tau - g', m_2) + \beta \left[\begin{array}{c} \Pi(g')\,\mathbb{E}_{m_1'} W_D(g', m_1') + \mathbb{E}_{A'} A' | (D \text{ party win}) + \\ [1 - \Pi(g')]\,\mathbb{E}_{m_2'} Z_D(g', m_2') \end{array} \right] \tag{9}$$

$$\text{s.t. } g' = G_R(g, m_2).$$

Recognizing that $Z_R(g, m_1, e, A) = Z_R(g, m_1) + e + A$ and proceeding as above yields the following recursions:

$$W_R(g, m_2) = \alpha U(g') + U(\tau - g', m_2) + \beta \left[\begin{array}{c} \Pi(g')\,\mathbb{E}_{m_1'} Z_R(g', m_1') + \mathbb{E}_{A'} A' | (D \text{ party win}) + \\ [1 - \Pi(g')]\,\mathbb{E}_{m_2'} W_R(g', m_2') \end{array} \right] \tag{10}$$

$$\text{s.t. } g' = G_R(g, m_2),$$

and

$$Z_R(g, m_1) = \alpha U(g', m_1) + U(\tau - g') + \beta \left[\begin{array}{c} \Pi(g')\,\mathbb{E}_{m_1'} Z_R(g', m_1') + \mathbb{E}_{A'} A' | (D \text{ party win}) + \\ [1 - \Pi(g')]\,\mathbb{E}_{m_2'} W_R(g', m_2') \end{array} \right] \tag{11}$$

$$\text{s.t. } g' = G_D(g, m_1).$$

Observe that while W_k and Z_k are, by definition, independent of e and A, they are dependent on *future* values of A' through the term $\mathbb{E}_{A'}|(D \text{ party win})$. This term recognizes that the election of the D party next period is not independent of the realized value of A'. In particular, as we show later in this section, the D party can win only if A' is above some threshold value that depends on g'. Thus, conditional on a D party win, the expectation of A' is nonzero.[20]

At the time of an election, voters compute their individual net gain from voting for each party and vote for the party for which this net gain is nonnegative. Since the value of m_1 or m_2 is realized after the election, the individual net gain to a type D person from voting for the D party is $\mathbb{E}_{m_1} W_D(g, m_1) + e + A - \mathbb{E}_{m_2} Z_D(g, m_2)$, and the individual net gain for a type R person from voting for the D party is $\mathbb{E}_{m_1} Z_R(g, m_1) + e + A - \mathbb{E}_{m_2} W_R(g, m_2)$. Given the pair (g, A), these expressions determine thresholds for the idiosyncratic shock, $e_k(g, A)$, $k \in \{D, R\}$, above which a k type will vote for the D party in the election. Specifically,

$$e_D(g, A) = -[\mathbb{E}_{m_1} W_D(g, m_1) - \mathbb{E}_{m_2} Z_D(g, m_2)] - A \tag{12}$$

and

$$e_R(g, A) = [\mathbb{E}_{m_2} W_R(g, m_2) - \mathbb{E}_{m_1} Z_R(g, m_1)] - A. \tag{13}$$

In these threshold expressions, the terms in square brackets represent the expected net gain to a person from his or her own party coming into power, ignoring the shocks e and A. Holding fixed A, the bigger the expected net gain term in (12), the lower the threshold e_D and the bigger the expected net gain in (13), the higher the threshold e_R. Hence, the larger these expected net gain terms, the more likely it is that an individual will vote for her own party. In contrast, an increase in A lowers *both* thresholds and increases the likelihood of all individuals voting for the D party.

Theorem 1. *Given W_k and Z_k, $k \in \{D, R\}$, the probability of the D party winning the election given g, $\Pi(g)$, is the probability that $A > A(g)$ where*

$$A(g) = \frac{1}{2} \left\{ [\mathbb{E}_{m_2} W_R(g, m_2) - \mathbb{E}_{m_1} Z_R(g, m_1)] - [\mathbb{E}_{m_1} W_D(g, m_1) - \mathbb{E}_{m_2} Z_D(g, m_2)] \right\}. \tag{14}$$

[20] There is no corresponding term for e' because the realization of an individual's e' has no consequence for whether or not the D party wins the election next period. Hence, the expectation of e' conditional on D party win is the unconditional expectation, which is 0.

Proof. By definition of $e_k(g, A)$, $k \in \{D, R\}$, the probability that a randomly selected voter will cast her vote in favor of the R party is

$$(1/2)F(e_D(g, A)) + (1/2)F(e_R(g, A)). \tag{15}$$

In what follows, it is assumed that this expression also gives the *fraction* of voters who vote for the R party, given (g, A).[21] Given that e has unbounded support, for any A, the expression in (15) is always in $(0, 1)$ and it is strictly decreasing in A. Hence, there is a unique A, denoted $A(g)$, such that exactly half of the people vote for the D party. For $A > A(g)$, more than half of the people will vote for the D party, and for $A < A(g)$, less than half will. Therefore, the probability that the D party wins the election is $\Pr[A > A(g)]$.

To derive the expression for $A(g)$ note that by definition

$$\frac{1}{2}F(e_D(g, A(g))) + \frac{1}{2}F(e_R(g, A(g))) = \frac{1}{2},$$

or

$$F(e_D(g, A(g)) = 1 - F(e_R(g, A(g)). \tag{16}$$

Since e is symmetrically distributed around 0, we may infer that at $A(g)$, e_D and e_R must be of opposite signs and equal in magnitude, i.e., $e_D(g, A(g)) + e_R(g, A(g)) = 0$. Using (12) and (13) then gives:

$$A(g) = \frac{1}{2}\left\{[\mathbb{E}_{m_2} W_R(g, m_2) - \mathbb{E}_{m_1} Z_R(g, m_1)] - [\mathbb{E}_{m_1} W_D(g, m_1) - \mathbb{E}_{m_2} Z_D(g, m_2)]\right\}.$$

□

The expression for $A(g)$ has intuitive properties. Observe that if the two expected net gain terms in square brackets are equal in value then $A(g) = 0$: When the net gain terms are equal, the fractions of people of either type who *cross vote* are also equal and the value of A needed to exactly balance the votes is 0. In contrast, if the net gain for D types is, say, smaller than the net gain for R types, then $A(g) > 0$: In this case, at $A = 0$, the fraction of D types cross voting will

[21] It is well understood that there is no "law of large numbers" that ensures this identification of probabilities as fractions when we are dealing with a continuum of voters (Judd (1985)). However, for our application, it is fine to simply assume that the "law" holds (see Feldman and Gilles (1985) and Uhlig (1996)).

be *larger* than the fraction of R types cross voting: D types don't benefit as much from having their party choose the composition of the public goods and so are more likely to be swayed by their idiosyncratic desire for or abhorrence of the *other* policies of the D party. Hence, a strictly positive value of A is needed to equate the probability of either party winning the election.[22]

The equilibrium of the model is defined as follows:

Definition 1. *A pure strategy Markov Perfect Equilibrium (MPE) is a collection of party value and policy functions* V_k^*, X_k^*, G_k^*, *a collection of voter value functions* W_k^*, Z_k^*, *and a pair of functions* $\Pi^*(g)$ *and* $A^*(g)$ *such that:*

- *Given* $G_R^*(g, m_2)$ *and* $\Pi^*(g)$, *the functions* $V_D^*(g, m_1)$ *and* $X_D^*(g, m_2)$ *solve* (1) - (2) *and* $G_D^*(g, m_1)$ *attains* $V_D^*(g, m_1)$
- *Given* $G_D^*(g, m_1)$ *and* $\Pi^*(g)$, *the functions* $V_R^*(g, m_2)$ *and* $X_R^*(g, m_1)$ *solve* (3) - (4) *and* $G_R^*(g, m_2)$ *attains* $V_R^*(g, m_2)$
- *Given* $G_D^*(g, m_1)$, $G_R^*(g, m_2)$, $\Pi^*(g)$, *and* $A^*(g)$, *the functions* $W_D^*(g, m_1)$ *and* $Z_D^*(g, m_2)$ *solve* (8) - (9)
- *Given* $G_D^*(g, m_1)$, $G*_R(g, m_2)$, $\Pi^*(g)$, *and* $A^*(g)$, *the functions* $W_R^*(g, m_2)$ *and* $Z_R^*(g, m_1)$ *solve* (10) - (11)
- *Given* $W_D^*(g, m_1)$, $Z_D^*(g, m_2)$, $W_R^*(g, m_2)$, *and* $Z_R^*(g, m_1)$, *the function* $A^*(g)$ *solves* (14) *and the function* $\Pi^*(g) = Pr[A > A^*(g)]$

For a model to be amenable to computation, it is important that there be easily verifiable conditions on model primitives for which the existence of at least one pure strategy MPE is assured. If this is the case, and the conditions hold for the model, one can be certain that a failure of an algorithm to find an equilibrium is a failure of the algorithm and not the result of a lack of internal consistency of the model. With this in mind, Appendix B gives sufficient conditions on model primitives for which the existence of a pure strategy is MPE assured. The conditions cover not

[22] In light of the key role of cross-voting, the formula for $A(g)$ is counterintuitive in one respect: It does not depend on the *specifics* of the CDF $F(e)$. To arrive at the recursions (8)-(11) that determine $\{W_j, Z_j\}$, the assumptions needed on e were that e is independent of all other shocks and that $\mathbb{E}e = 0$. And the assumption needed on e to go from (15) to the expression for $A(g)$ is that e is has unbounded support and is distributed symmetrically around zero. Aside from these properties, the precise shape of $F(e)$ does not matter for the determination of $A(g)$. This somewhat paradoxical result stems from the fact that there are equal measures of the two types of people. If this were not true, (16) would not be true and $A(g)$ *would* depend on the specific shape of $F(e)$.

only the model described so far but also variants discussed later in the paper and, potentially, other models featuring endogenous election probability. We have:

Theorem 2. *Under Assumptions 1 – 6 stated in Appendix B, a pure strategy MPE exists.*

Proof. See Appendix B.

For the model described thus far, the key requirements are that g and g' belong to a finite set (i.e., the interval $[0, \tau]$ be replaced by a discrete approximation), m_ℓ have compact support, the CDFs for m_ℓ, A and e be continuous, and for any $x \neq \hat{x}$, the difference $U(x, m_\ell) - U(\hat{x}, m_\ell)$, $\ell = \{1, 2\}$, be strictly increasing or strictly decreasing in m_ℓ.

4 Policy Inertia and Incumbency Disadvantage of Parties in a Two-Period Model

The main result of this paper — incumbency disadvantage — can be illustrated and explained in a two-period setting. Let $t = 1, 2$. Let g' and g'' denote periods 1 and 2 policies, respectively. Since $t = 1$ is the initial period, there are no adjustment costs associated with the choice of g'. For period 2, it is assumed that $\psi(g'', g') = \eta(g'' - g')^2$, $\eta > 0$. In this two-period setting, there is no need for the m_ℓ shocks and so they are dropped. For analytical tractability $\alpha = 0$ and $U(x) = -(\tau - x)^2$. Thus, the period 2 utility of D and R types from the composition $(g'', \tau - g'')$ is $U(g'') = -(\tau - g'')^2$ and $U(\tau - g'') = -(\tau - (\tau - g''))^2 = -g''^2$, respectively. In terms of g'', the ideal policy of D types is τ and the ideal policy of R types is 0.

To solve for the equilibrium of this model by backward induction, we will first show that if the inherited policy in period 2 is closer to what the D types (R types) prefer, the D party (R party) is less likely to be elected in period 2.

Consider the policy of the D party if it wins the election in period 2. It chooses g'' to maximize $-(\tau - g'')^2 - \eta(g' - g'')^2$ subject to $g'' \in [0, \tau]$. This maximization implies

$$G_D(g') = \frac{1}{1+\eta}\tau + \frac{\eta}{1+\eta}g' .$$

The optimal decision is, thus, a convex combination of the D type's ideal policy, τ, and the inherited policy g'. Similarly, if the R party wins in period 2,

$$G_R(g') = \frac{\eta}{1+\eta} g'.$$

This decision is a convex combination of the R type's ideal policy, 0, and the inherited policy g'.

Given the parties' postelection choices in period 2, the net gain to a D type from having the D party choose policies in period 2 is

$$-\left[\frac{\eta}{1+\eta}\right]^2 (\tau - g')^2 + e' + A' + \left(\tau - \frac{\eta}{1+\eta} g'\right)^2$$

and the net gain to an R type from having the R party choose policies in period 2 is

$$-\left[\frac{\eta}{1+\eta}\right]^2 g'^2 + \left(\tau - \frac{\eta}{1+\eta}(\tau - g')\right)^2 - A' - e'.$$

Then the value of e such that a realization of e' above that value will make a D type vote for the D party is

$$e_D(g', A') = \left[\frac{\eta}{1+\eta}\right]^2 (\tau - g')^2 - \left(\tau - \frac{\eta}{1+\eta} g'\right)^2 - A', \tag{17}$$

and the value of e such that a realization of e' above that value will make an R type also vote for the D party is

$$e_R(g', A') = -\left[\frac{\eta}{1+\eta}\right]^2 g'^2 + \left(\tau - \frac{\eta}{1+\eta}(\tau - g')\right)^2 - A'. \tag{18}$$

Then, as shown in the previous section, the value of A such that a realization of A' above that value will result in a D party win solves the equation $e_D(g', A') + e_R(g', A') = 0$. Using (17) and (18), this threshold value of A' is given by the following simple expression:

$$A(g') = \frac{2\eta\tau}{(1+\eta)^2}\left(g' - \frac{\tau}{2}\right). \tag{19}$$

When $\eta > 0$, the sign of $A(g)$ depends on the sign of $g - \tau/2$. If g is closer to the ideal choice of the D party (R party), then $A(g)$ is positive (negative), which means that the probability of the D party (R party) winning the election is *less* than one-half.

To understand this result, we can examine which party wins the election when A is equal to zero. The magnitude of the net gain for each type from having his own party choose policies (ignoring the voter preference shocks e) is

$$|U_k(G_D(g')) - U_k(G_R(g'))| \approx \quad |U_k'\left(\tfrac{G_D(g')+G_R(g')}{2}\right)| \cdot |G_D(g') - G_R(g')|. \tag{20}$$

The type for which this magnitude is larger will see a higher share voting for its own party as fewer of them will be swayed by the idiosyncratic shock and, thus, will win the election. The magnitude will be higher for the type for which $|U_k'\left(\frac{G_D(g')+G_R(g')}{2}\right)|$ is higher. Given that the midpoint of policies desired by the two parties is $\frac{\tau}{2} + \frac{\eta}{1+\eta}\left(g' - \frac{\tau}{2}\right)$, by diminishing marginal utility $|U_D'\left(\frac{G_D(g')+G_R(g')}{2}\right)|$ will be smaller (larger) than $|U_R'\left(\frac{G_D(g')+G_R(g')}{2}\right)|$ for g' greater than (smaller than) $\tau/2$, and so, the R party (D party) will win the election when $A = 0$.

Note that $A(g) = 0$ for all g if $\eta = 0$ or if $\eta = \infty$. If $\eta = 0$, there are no adjustment costs and the winning party implements its ideal policy regardless of g. Then the midpoint of the desired policies of the two parties is $\tau/2$ and $U_D'(\tau/2)$ is equal to $U_R'(\tau/2)$, and the magnitude of the net gain from voting one's party is the same for both types. If $\eta = \infty$, adjustment costs are infinite so $g' = g$ regardless of which party is elected. In this case, the $U_D'(g')$ will be different from $U_R'(g')$ depending on g, but this difference does not matter because $|G_D(g') - G_R(g')| = 0$ (see equation (20)).

Overall, the results for $0 < \eta < \infty$ will be consistent with an incumbency disadvantage if the D party chooses $g' > \tau/2$ in period 1 when in power (and, similarly, if the R party chooses $g' < \tau/2$ in period 1 when it is in power). To solve the period 1 decisions of the parties analytically, we will assume that A is distributed uniformly with support $[-\bar{A}, \bar{A}]$. Then, $\Pi(g')$, the probability that the D party wins the election in period 2, takes the following simple form:

$$\Pi(g') \equiv \Pr[A \geq A(g')] = \begin{cases} 0 & \text{if } A(g') \geq \bar{A} \\ \left[\frac{1}{2} - \frac{\eta\tau}{(1+\eta)^2 \bar{A}}\left(g' - \frac{\tau}{2}\right)\right] & \text{if } \bar{A} > A(g') > -\bar{A} \\ 1 & \text{if } -\bar{A} \geq A(g'). \end{cases} \tag{21}$$

We also assume that $\Pi(g')$ is in $(0,1)$ for all $g' \in [0,\tau]$; i.e., for all feasible choices of g', the probability of any party winning the election next period is strictly positive. From (21), it can be verified that this will be the case if

$$\bar{A} > \frac{\eta}{(1+\eta)^2}\tau^2, \tag{22}$$

i.e., if there is sufficient uncertainty about the aggregate voter preference shock. For convenience, denote the r.h.s. of (22) as ϕ. Then, the assumption is that $\phi/\bar{A} < 1$.

Given these postelection choices, the payoffs to the D party from winning and losing the election in period 2 are, respectively,

$$W_D(g') = -\left[\frac{\eta}{1+\eta}\right](\tau - g')^2 \text{ and } X_D(g') = -\left(\frac{1}{1+\eta}\right)\tau^2 - \left(\frac{\eta}{1+\eta}\right)(\tau - g')^2,$$

and the payoffs to the R party from winning and losing the election in period 2 are, respectively,

$$W_R(g') = -\left[\frac{\eta}{1+\eta}\right]g'^2 \text{ and } X_R(g') = -\left(\frac{1}{1+\eta}\right)\tau^2 - \left(\frac{\eta}{1+\eta}\right)g'^2.$$

Given these payoffs, the period 1 decision problem of the D party when it is in power is

$$\max_{g' \in [0,\tau]} \quad -\left(1 + \frac{\beta\eta}{1+\eta}\right)(\tau - g')^2 - \beta[1 - \Pi(g')]\left(\frac{1}{1+\eta}\right)\tau^2, \tag{23}$$

where $\Pi(g')$ is the probability of the D party winning the election in period 2. Similarly, the period 1 decision problem of the R party when it is in power is:

$$\max_{g' \in [0,\tau]} \quad -\left(1 + \frac{\beta\eta}{1+\eta}\right)g'^2 - \beta\Pi(g')\left(\frac{1}{1+\eta}\right)\tau^2. \tag{24}$$

Turning first to the D party, it follows from (23) and (21) that the net marginal gain to the D party from increasing g' is proportional to

$$2(\tau - g')(1 + \eta + \beta\eta) - \beta\left(\frac{\phi}{\bar{A}}\right)\tau. \tag{25}$$

This net marginal gain is strictly negative for $g' = \tau$, since both β and $\phi/\bar{A}$ are strictly positive and it is strictly positive at $g' = 0$ because both β and $\phi/\bar{A}$ are less than unity. It follows that the

D party's optimal choice of g' is in the interior of $[0, \tau]$. Then, the D party's period 1 choice of g', g'_D, must satisfy

$$(\tau - g'_D) = \frac{\beta}{1 + \eta + \beta\eta} \left(\frac{\phi}{\bar{A}}\right) \frac{\tau}{2}. \tag{26}$$

Since the term multiplying $\tau/2$ on the r.h.s. is strictly less than 1, $g'_D > \tau/2$. Then, by (21), the D party's probability of reelection in period 2 will be less than one-half. Thus, the D party will be at an electoral disadvantage going into the elections in period 2. By symmetry of the model, the R party's optimal choice of g' in period 1 is less than $\tau/2$ and by (21) again, the R party will be at an electoral disadvantage going into the election in period 2.

We can summarize these results in:

Proposition 1. *If the distribution of A is uniform and the probability of the election of either party is strictly positive for all feasible choices of g', the optimal choice of g' in period 1 by either party implies that the incumbent party is disadvantaged in the period 2 election.*

One additional point can be made with this simple model. Note that the assumption that only *parties* bear the costs of adjusting policies is important to the conclusion of Proposition 1. If the costs of changing policies also entered the payoff functions of the two types of people, then, ignoring the voter preference shocks, the net gain to type k from having the k party choose policies will be given by $W_k(g') - X_k(g')$, $k \in \{D, R\}$. But this difference is $[1/(1+\eta)]\tau^2$, which is independent of k and g' and so $\Pi(g') = 0.5$ for all g'. Therefore, there will be no incumbency disadvantage or advantage going into the elections in period 2. However, this property will not necessarily be true in the other models of inertia we analyze in the quantitative section.

Before closing this section, we clarify why we did not need the m_ℓ shocks for equilibrium in the two period model, but these shocks are essential in the infinite horizon model. If $[0, \tau]$ is discretized, one can always compute, via backward induction, the equilibrium decision rules $G_k(g, n)$ — where n is the number of periods to the terminal period — for *any* $n \in \mathbb{N}$. The problem, however, is that $|G_k(g, n) - G_k(g, n+1)|$ may fail to vanish for all g as $n \to \infty$. In our experience, this problem arises when there is time inconsistency, which leads to nonconcavity of the continuation value functions.[23]

[23] For large n, changes in continuation values can be small as n is incremented, but, even small changes can induce jumps in $G_k(g, n+1)$ because of the nonconvexities in the decision problems of the parties. As a result, neither decision rules nor continuation value functions settle down to stationary functions as n is increased.

5 Incumbency Disadvantage in a Quantitative Dynamic Model

In this section, our goal is to show not only that the incumbency disadvantage can arise in the dynamic model but also that the model is capable of delivering the observed *magnitude* of this effect. Given the quantitative nature of our goal, the demonstration proceeds by choosing a realistic parameterization of the model and examining the relevant properties of the computed MPE. As part of our goal to understand why incumbents act the way they do, we highlight the key roles played by policy inertia and uncertainty in election outcomes in generating incumbency disadvantage.

To proceed with the quantitative analysis, we adopt some parametric assumptions. We assume that $U(x, m_\ell) = (x+m_\ell)^{1-\gamma}/[1-\gamma]$, $\ell = 1, 2$. For the base model, we assume the $\psi(g, g') = \eta(g-g')^2$, $\eta > 0$. The idiosyncratic shock $e \sim N(0, \sigma_e^2)$. The aggregate shock $A \sim \text{Unif}([-\bar{A}, \bar{A}])$ and the distributions of the party preference shocks m_ℓ, $\ell = 1, 2$, are both $\text{Unif}([-\bar{m}, \bar{m}])$. Note that both A and the m_ℓ distributions are symmetric around 0.

Turning to parameter values, since national elections happen every two years, the value of β is set to 0.92, which corresponds to a biennial discount rate of 8 percent. The value of γ is set to 2. The value of τ is normalized to 1.[24] Since we don't observe large shifts in expenditure patterns when parties controlling the presidency change, α is set conservatively to 0.90 — this implies that voters' static optimum is to have their party spend 51.3 percent of the total budget on their preferred good.

The remaining parameters, namely η and the dispersions of the distributions of A, e and m_ℓ, have important effects on the magnitude of the incumbency disadvantage in the model. Of these, the dispersion of m_ℓ is special in that having a large enough dispersion helps us compute an equilibrium.

Conditional on a dispersion of m_ℓ and the parameters listed earlier, experimentation showed that the model can deliver the observed magnitude of the incumbency disadvantage for a *range* of values of the remaining three parameters. Among the many constellations of parameter values we could pick to be consistent with the observed magnitude of the incumbency disadvantage, we chose the one in which η is 4.8, the standard deviation of e is 0.02, and the support of (the uniformly distributed) A is ± 0.01.

[24] But we restrict the feasible set of g' to be $(\bar{m}, 1-\bar{m})$. Then $g' + m_1$ and $\tau - g' + m_2$ are strictly positive for any choice of g' and any realization of m_ℓ.

These parameter choices are summarized in Table 4.

Table 4: Parameter Selections

Parameter	Description	Value
γ	Curvature of utility function	2.00
β	Biennial discount factor	0.92
τ	Total government exp.	1.00
α	Weight given to other party's desired public good	0.90
$\overline{m}$	Support of party preference shock, $\pm\overline{m}$	0.01
η	Adjustment cost parameter	4.8
σ_e	S.D. of idiosyncratic voter preference shock	0.02
$\overline{A}$	Support of aggregate voter preference shock, $\pm\overline{A}$	0.01

To confirm that the model generates the incumbency disadvantage, Table 5 reports the results of regressions run on model-generated data that mimic the regressions reported in the top panel of Table 2. The dependent variable is the D party's vote share and the explanatory variables are trichotomous variables that take on the values $+1$, -1, or 0 depending on whether the D party has been in power for three or more (or four or more) model periods, or the R party has been in power for three or more (or four or more) model periods, or neither. For comparison purposes, we also record the outcome of the same regression with the probability of a D victory as the dependent variable.

Table 5: Incumbency Disadvantage

		Model		
Dep. Var.	$\%D$	$\%D$	Prob. of D Win	Prob. of D Win
Constant	50.00	50.00	0.50	0.50
SIX^+	-2.42	-	-0.07	-
$EIGHT^+$	-	-2.44	-	-0.07

The first two columns in Table 5 report the magnitude of the incumbency disadvantage in the model for both six or more years and eight or more years of incumbency (in the model, these correspond to three or more periods or four or more periods). For either measure of incumbency, the incumbency disadvantage is 2.4 percentage points. The next two columns report the incumbency disadvantage in terms of the decline in the probability of reelection. For either measure of incumbency, the decline in the reelection probability is 7 percentage points. We also confirmed that adjustment costs remain central to the incumbency disadvantage in the full dynamic setting:

If we set $\eta = 0$ and run the same regressions on the model output as in Table 5, the coefficients on the incumbency variables are all estimated to be zero.

5.1 Role of Policy Inertia

A positive η introduces inertia in policies and creates a dynamic link between periods. A consequence is that a party's long-run ideal policy, namely the average composition of spending toward which it tends as its incumbency lengthens, deviates from its no-inertia ideal policy, i.e., its average policy when $\eta = 0$.

Figure 1 charts, for different values of η, the relationship between the average expenditure (over a long simulation) on the preferred good of the incumbent party against the party's years of incumbency. The blue dotted line is the long-run ideal policy, corresponding to the $\eta = 0$ case. As the line shows, the party immediately goes to its long-run ideal and the line is flat at 0.513. The solid black line immediately below corresponds to our base model with $\eta = 4.8$. For the base model, the average expenditure is initially below its long-run level of 0.512 but reaches that level by the sixth year of incumbency and stays flat thereafter. As η is increased, the long-run expenditure level shifts down and the years of incumbency needed to converge to the ideal level lengthens. As seen in the shape of the red dashed line ($\eta = 10$), policies start out closer to 0.50 and continue to move up even beyond the eighth year of incumbency.

These expenditure dynamics have implications for the time path of the incumbency disadvantage. Figure 2 plots the average percentage of voters who cast their ballots in favor of the incumbent. There is no incumbency disadvantage when $\eta = 0$ and the dotted blue line is flat at zero. For the base model, the share falls by about 1.5 percentage points at the end of two years of incumbency and the disadvantage continues to increase until the share stabilizes at around 47.5 percent by the sixth year of incumbency. The incumbency disadvantage increases with η, as shown in the red dashed line corresponding to $\eta = 10$. But as η rises enough, the incumbency disadvantage weakens and eventually disappears when $\eta = \infty$. The nonmonotonic relationship between the strength of the incumbency disadvantage and η was explained in the context of the two-period model and follows the same logic in the fully dynamic model: When the cost of adjustment is high enough, neither party can change policies too much and, consequently, the incumbency disadvantage weakens.

Figure 1:
Incumbency and Average Expenditure on the Preferred Good

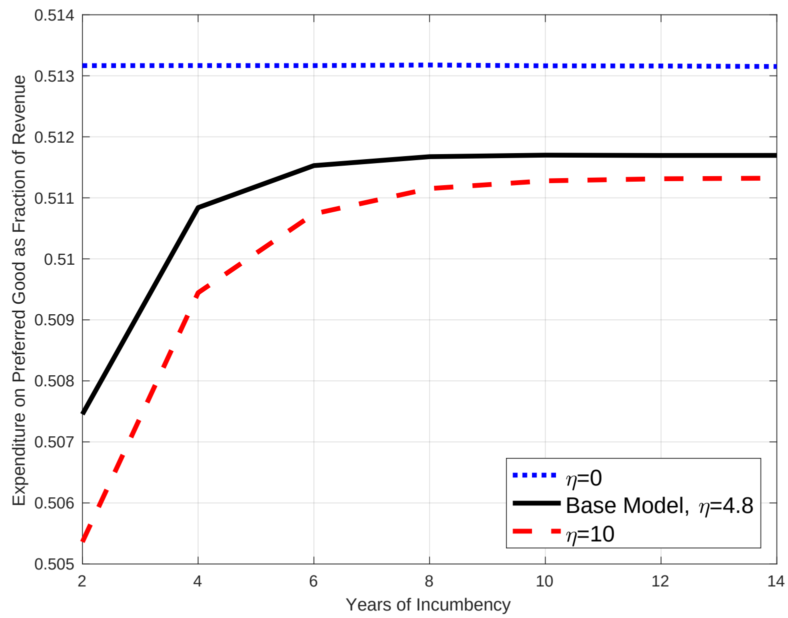
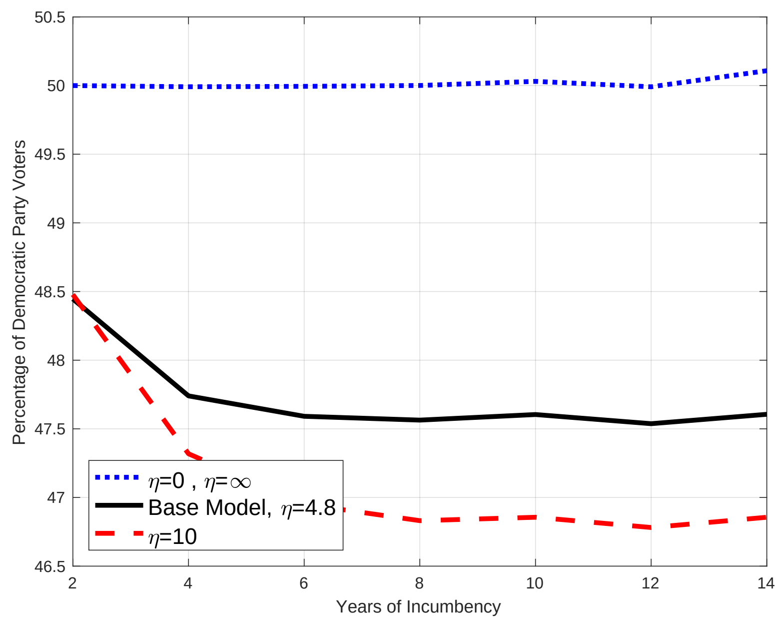


Figure 2:
Incumbency and Average Lead in Elections

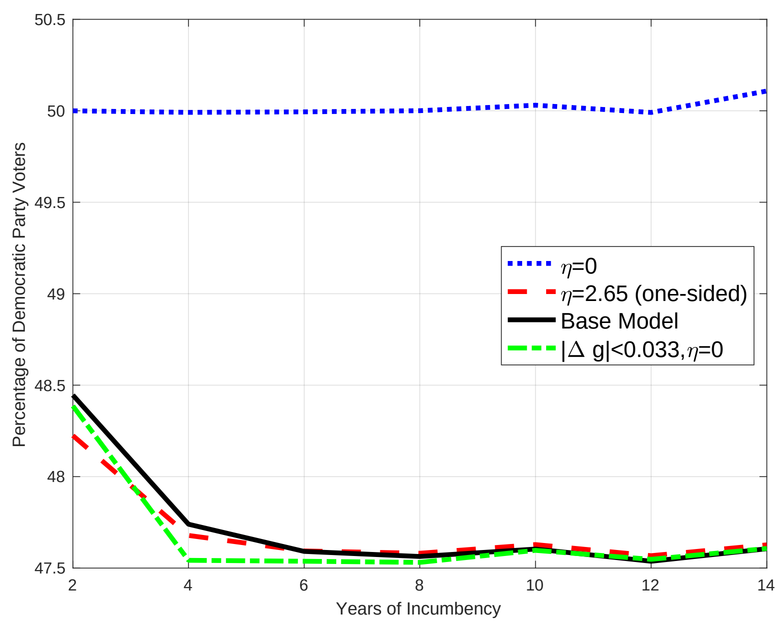


5.2 Alternative Models of Inertia

In this section, we show that the incumbency disadvantage also occurs for other models of inertia that may seem plausible. For this, we examine two alternative models. In the first alternative, there is a quadratic cost of adjustment just as in the base model, but the costs of adjustment are borne only by the party in power. In the second alternative, there is no cost of adjustment, but there is an upper bound Δ on how much policies can change in either direction in any period (we call this the constraint-on-change model). In these experiments, all parameters unrelated to η or Δ are the same as in the base model. We pin down the new adjustment cost parameters — the one-sided η and Δ — to match the same SIX^+ coefficient as in the base model. The value of η is now 2.65, and the value of Δ is 0.033.

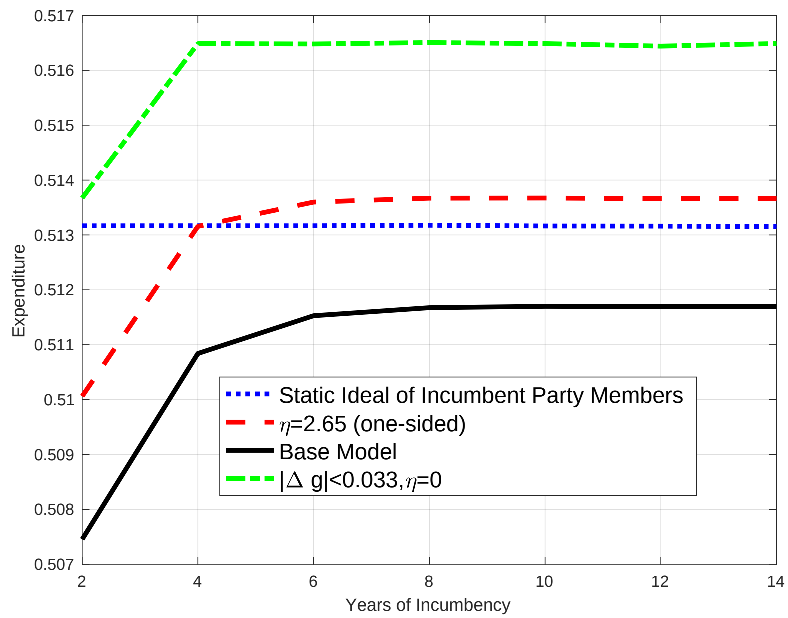
The first point is that the pattern of the incumbency disadvantage documented in Tables 1–3 can be accounted for by any of these alternative models of inertia. Figure 3 plots the average vote share in elections for the two alternative models along with the base model. The predicted relationships are virtually identical.

Figure 3: Alternative Models of Inertia
Incumbency and Average Lead in Elections



Second, the alternative models imply quite different long-run ideal policies of each party. In contrast to the base model, the long-run ideal policies are *more* extreme than the static ideal policy

Figure 4: Alternative Models of Inertia Incumbency
and Expenditure on the Preferred Good



of their party members, as shown in Figure 4. Because of inertia, the incumbent party pushes beyond the static ideal policy of its party members to assure them policies "close" to their static ideal even when it is out of power and the other party is choosing policies. Thus, inertia can explain why parties throw their weight behind extreme policies. Interestingly, this extremism does not arise in the base model because the costs of changing policies are borne by *all* representatives in government. Since an incumbent party anticipates the swing back in policy in future periods — and the costs associated with that reversal — it becomes more circumspect about pushing for policies that depart too far from the ideal policies of the other party.

These differences in expenditure patterns across the different models of inertia also mean that the implied welfare of voters differs across the models. From an ex-ante perspective, a more volatile expenditure pattern between the two public goods is costly because of diminishing marginal utility: In some periods, the marginal utility from the preferred good of some type is high, and in other periods, it is low. As mentioned in the introduction, the existing literature has noted that such cycles can be welfare reducing.

Table 6 reports the identical lifetime utility of both types of voters when the inherited policy is $g = 0.50$, and so there is a 50 percent chance of either party being elected. The findings confirm our intuition: Welfare is highest for the base model in which the standard deviation of g is lowest and welfare is lowest for the constraint-on-change model for which the standard deviation of g is the highest.

Table 6: Welfare Implications of Policy Inertia

Models	Welfare Loss in % Relative to Base Model	Std. Dev of g'
Base model	-	0.0095
$\eta = 2.65$ (one-sided)	0.02	0.0119
$\lvert\Delta g\rvert < 0.033$	0.07	0.0158

6 Policy Inertia and Downsian Logic Without Preelection Commitment

Downs (1957) famously observed that if politicians care only about winning elections, the policies enacted will converge to the centrist ones. Alesina (1988) pointed out that if parties also care about the policies they enact and cannot commit to policies before an election, then equilibrium policy will be the one that the winning party most prefers, *regardless* of how motivated the party is to win the election. In other words, the inability to commit can completely short circuit the moderating influence of office motivation.

Here we show that inertia and prospective voting can reintroduce the Downsian logic even when there is no preelection commitment to policies. For this, we extend the base model to include a utility benefit enjoyed by the political party when it is in power. Specifically, the period utility from being in power for the D party is $U(g', m_1)+\alpha U(\tau-g')-\eta(g'-g)^2+B$, where $B > 0$. Symmetrically, the period utility from being in power for the R party is $\alpha U(g') + U(\tau - g', m_2) - \eta(g' - g)^2 + B$. The constant B is a stand-in for the office motivation of the representatives seeking election.

We confirm Alesina's result for the no-inertia version of this model.

Theorem 3 (Alesina 1988)**.** *If there are no adjustment costs or constraints on changing policies, parties choose their statically ideal policies regardless of the value of B.*

Proof. If $\eta = 0$, g is no longer payoff-relevant and so its value cannot affect equilibrium outcomes. Thus, $A^*(g)$ and $G_k^*(g, m_\ell)$, $(j, \ell) \in \{(D, 1), (R, 2)\}$ are independent of g but, potentially, dependent on B. Assume that the former is $A^*(B)$ and the latter are $G_k^*(m_\ell, B)$. Then, the continuation value of party D when it is in power is

$$\Pr[A \geq A^*(B)]\{\mathbb{E}_{m_1} V_D^*(m_1', B) + \mathbb{E}_A[A|A \geq A^*]\} + [1 - \Pr[A < A^*]]\{\mathbb{E}_{m_2} X_D^*(m_2', B)\}.$$

Since this continuation value is independent of g, once party D is elected, the best it can do is solve its static optimization problem. Thus,

$$G_D^*(m_1, B) = \text{argmax}_{g' \in [0,\tau]} U(g', m_1) + \alpha U(\tau - g') + B,$$

which is evidently independent of B. Symmetrically, the R party will choose its statically ideal policy, independent of the value of B. As a corollary, one may verify that when the two parties act in this way, the net gain terms within square brackets in (14) are equal, and so, $A^*(B) \equiv 0$ and the probability of reelection is one-half regardless of B. $\square$

This result changes when $\eta > 0$. With inertia, the presence of B creates a conflict of interest between representatives and their constituencies: Upon electoral defeat, the party loses B and the prospect of this loss restrains the party from pushing as hard for its constituents as it otherwise would. We confirm this in Figure 5. The top panel plots the average equilibrium value of g' against the number of years of incumbency for the baseline model and the baseline model extended with a $B = 0.04$. The graph shows that as the D party comes into power and continues in power, the average g' rises for both models. However, the rise in g' for the base model is more pronounced than in the model with office motivation. Figure 6 plots the average lead of the party in power as the incumbency progresses. The average lead falls in both cases, but the decline is less pronounced for the model with office motivation.

Interestingly, although office motivation creates a conflict of interest between the party and its constituents, the welfare implications of office motivation for the constituents are *positive*. This is because office motivation reduces the amplitude of the policy cycles discussed in the previous subsection. Table 7 reports the welfare effects of office motivation. As before, the welfare measure is the (identical) lifetime utility of types, conditional on $g = 0.50$. Observe that welfare is higher and the standard deviation of g is lower with office motivation than without.

Table 7: Welfare and Office Motivation

Models	Welfare Gain Relative to Base Model, in %	Std. Dev. of g'
Base Model	–	0.0095
Base Model with $B = 0.04$	0.02	0.0065

Figure 5:
Office Motivation, Incumbency, and Average Expenditure on the Preferred Good

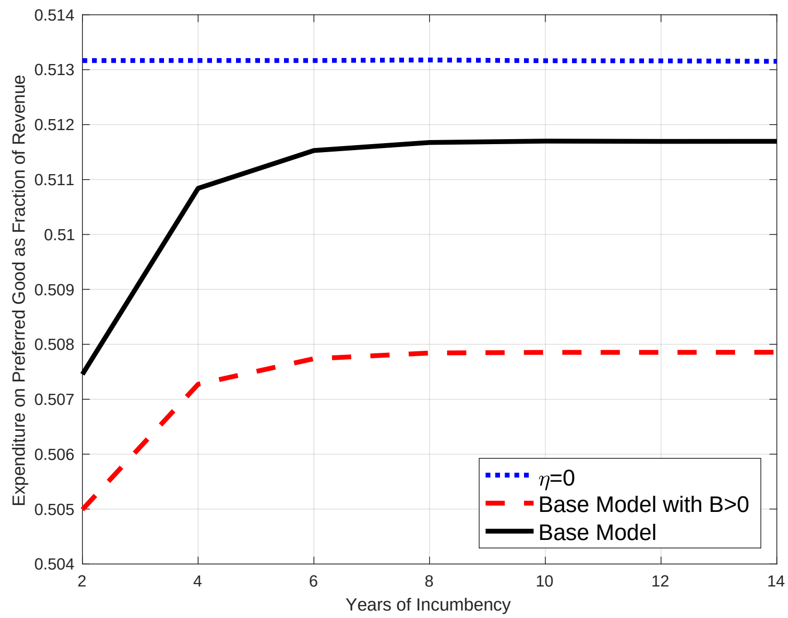
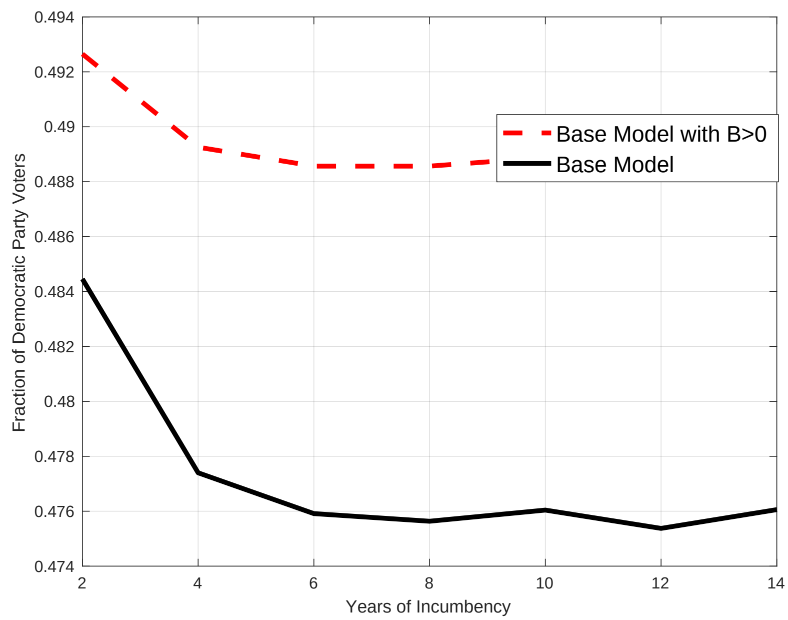


Figure 6:
Office Motivation, Incumbency, and Average Lead in Elections



7 Multiple Equilibria in Low-Noise Environments

Here we alert the reader to the fact that if $\bar{m}$ is low, the regular algorithm typically fails to converge to an equilibrium. In those few cases in which an equilibrium was found (this was for very low $\bar{A}$ values), the equilibrium found was *asymmetric*. Since asymmetric equilibria always come in pairs (with the roles of the two parties reversed) and since equilibria generically come in odd numbers, in those cases, there are at least two asymmetric equilibria and at least one symmetric equilibrium. The existence of the symmetric equilibrium was confirmed by forcing the algorithm to look only for symmetric equilibria (this is done by restricting the continuation values of the two parties to always be mirror images of each other). With this restriction, a symmetric equilibrium could be found even for low values of $\bar{m}$. However, the symmetric equilibria so found are unstable in the following sense: If the regular (unrestricted) algorithm starts off with the equilibrium continuation values found by *imposing* symmetry, the algorithm eventually diverges from the symmetric equilibrium.[25] This explains why we cannot find a symmetric equilibrium (without imposing symmetry) when $\bar{m}$ is low. In addition, our algorithm of value function iteration with slow update of future values typically fails to find asymmetric equilibria. For these reasons, throughout this study, the $\bar{m}$ was set high enough so that our algorithm could find equilibria without imposing symmetry.

For the interested reader, we describe one asymmetric equilibria we find when $\bar{A}$ is set to a low value of 0.00001 and $\bar{m}$ is 0.001. In this case, the equilibrium involves one of the parties staying in power forever. Focusing on the equilibrium in which the D party is in power forever, the equilibrium $\Pi(g) = 1$ for all $g < 0.5064$, i.e., $\Pi(g = 0.5) = 1$, which is never the case in any symmetric equilibrium (in a symmetric equilibrium $\Pi(g = 0.5) = 0.5$). In this equilibrium, the reason the D party is more attractive to voters around $g = 0.5$ is that the D party understands that if it chose $g' > 0.5064$, its probability of reelection will fall drastically (this is a feature of the volatility of A being very low). So it sticks with a moderate policy that is at or below 0.5064. In contrast, if the R party were to come into power, it faces no such disciplining device: The R party desires a policy that is less than 0.5 (which is less than 0.5064) and, for any such choice, the D party will win the election for sure. Given that it cannot increase its reelection probability by being more disciplined, it follows extreme policies. This behavior, in turn, makes people, in the aggregate, prefer the D party to the R party at $g = 0.50$.

[25] Numerically, any equilibrium found is approximate. When the regular algorithm starts off with the approximate symmetric equilibrium found by imposing symmetry, the small approximation error is enough to cause the algorithm to diverge.

If we hold $\bar{A}$ at 0.0001 but increase $\bar{m}$ to 0.04, a symmetric equilibrium is found. In this symmetric equilibrium, power changes infrequently: The incumbent party picks policies very close to 0.5 (in fact, the chosen policy is very slightly closer to the other party's ideal point) most of the time and when the m_ℓ is such that the marginal utility from its preferred good is very high, it picks an extreme policy and loses power. Once again, this behavior is a consequence of the volatility of A being very low. If we hold the value of $\bar{m}$ at 0.04 but raise the value of $\bar{A}$ to the base model value of 0.01, the equilibrium found is symmetric and this equilibrium displays the incumbency disadvantage property of the base model.

8 Summary

In this paper, we documented a strong pattern of incumbency punishment in U.S. politics. Postwar evidence of the electoral performance of the two parties in House and state gubernatorial elections show that a long incumbency of a party leads to substantial decline in the popularity of the party in elections. We used Alesina and Tabellini's (1990) model of partisan politics, extended to have elections with prospective voting, to explain this finding. We showed that the costs of changing policies, or simply constraints on how much policies can change from one period to the next, combined with uncertainty in election outcomes, can generate incumbency disadvantage.

We examined the implications of policy inertia for how parties choose policies. We showed that inertia can cause parties to target policies that are more extreme than the policies they would support in the absence of inertia and that such extremism can be welfare reducing. On the other hand, inertia implies that office motivation matters for policy choice, even when there is no preelection commitment to policies, and this can dampen policy cycles and raise welfare.

APPENDIX

A Empirical Appendix

Table 8 reports the results of U.S. House elections using the Democratic seat share won as the dependent variable. The SIX^+ coefficient is again significant and the magnitude is more than double the coefficient in the main regression. It appears that a given decline in vote share leads to a larger proportional loss in seat share.

Table 8:
Presidential Incumbency of a Party and Democratic Share of House Seats

DEP. VAR.	*DV*
CONSTANT	54.33*** (1.95)
SIX^+	−5.76*** (1.74)
$PDPIGR * TWO^+$	0.88** (0.54)
$MIDTERM * TWO^+$	−0.28 (1.87)
SD(DEP VAR)	7.27
NO. OF OBS.	32
R^2	0.39
ADJ. R^2	0.32

Table 9 reports the results for gubernatorial elections when the duration of incumbency is defined as two or more years. The incumbency coefficient of TWO^+ is again significant at the 1 percent level.

Table 9:
Gubernatorial Incumbency of a Party and Its Share of Gubernatorial Votes

DEP. VAR.	*GDV*		*GDVN*	
TWO^{+}	-2.24^{***}	-1.54^{***}	-4.62^{***}	-3.67^{***}
	(0.52)	(0.45)	(0.83)	(0.57)
$PPIGR \cdot TWO^{+}$	0.34	-0.03	.56	0.07
	(0.22)	(0.08)	(0.34)	(0.10)
INC		7.59^{***}		9.30^{***}
		(0.57)		(0.70)
YEAR FE	YES	YES	YES	YES
STATE FE	NO	NO	YES	YES
SD DEP. VAR.	6.19	9.22	9.10	10.40
NO. OF OBS	158	783	153	770
R^2	0.58	0.40	0.80	0.38
$ADJ.R^2$	0.33	0.34	0.52	0.27

B Existence of a Pure Strategy Markov Perfect Equilibrium

The goal of this Appendix is to provide a secure computational foundation for models with endogenous elections. The Markov Perfect Equilibria (MPE) of such models are challenging to compute because decision problems that arise in these models can, very naturally, give rise to nonconcave objective functions. Consequently, the MPE decision rules of such models need not be continuous in state variables and the value functions need not be everywhere differentiable in state variables. Computational methods that rely on the differentiability of value functions are, thus, ill suited for this class of models and such models are best solved on a grid.

But even if state and action spaces are finite, the nonconvexities inherent in the decision problems of the parties imply that a stationary MPE may call for randomization over actions that give the same (lifetime) utility to the party. But allowing for randomization in dynamic games with large state $\times$ action spaces (even when attention is restricted to subgame perfect equilibria) is also computationally challenging. The continuously distributed shock to primitives (the m_ℓ shocks in the main model) are introduced to obviate the need for randomization over actions.[26]

The main contribution of this Appendix is to give easily verifiable conditions on primitives for which at least one pure strategy MPE is assured and can be computed. These conditions apply to a class of models that includes all the models discussed in the main text and, potentially, other models of interest.

B.1 The Assumptions on Primitives

Finite States and Actions:

Let $\mathcal{I} = \{1, 2, \ldots, I\}$, $I \geq 2$ be the set of possible endogenous states the economy can be in at the start of any given period.[27] We use i and j to denote generic elements of $\mathcal{I}$. The action space when the state is i and party k is making decisions is denoted $\Gamma_i^k \subseteq \mathcal{I}$. We say $\{i, j, k\}$ is a *feasible triple* if $j \in \Gamma_i^k$.

Assumption 1. *$\Gamma_i^k \neq \emptyset$ for all $i \in \mathcal{I}$ and all $k \in \{D, R\}$.*

[26] This is similar to Harsanyi's (1973) purification of mixed strategies in normal form games via additive payoff perturbations.

[27] The fact that $\mathcal{I}$ contains only endogenous states is also not restrictive. The proof of existence can be straightforwardly extended to include any number of discrete shocks that affect feasible sets. The same is true for the computation method.

In the models with quadratic adjustment costs, $\Gamma_i^k = \mathcal{I}$ and is independent of i and k. In the "constraint-on-change" model, Γ_i^k is independent of k but not i. In the first case, Assumption 1 is satisfied by virtue of $\mathcal{I}$ being nonempty; in the second case, it is satisfed because $i \in \Gamma_i^k$.

Current period rewards:

Let $u_{i\,j}^k(m_k)$ denote the current period reward to a k type if party k is in power, the state is i, the preference shock is m_k and j is chosen, and let $\tilde{u}_{i\,j}^k(m_{\sim k})$ denote the current period reward to a k type if party $\sim k$ is in power, the state is i, the preference shock is $m_{\sim k}$ and j is chosen. Here $m_k \in [\underline{m}, \overline{m}] \equiv M \subset \mathbb{R}$.[28]

When the party in power is D, the utility flows to all individuals are augmented by $e + A$, where $e \in \mathbb{R}$ and $A \in \mathbb{R}$ are the idiosyncratic and aggregate components of an individual's net preference for the D party.

Let $U_{i,\,j}^k(m_k)$ denote the current period reward to party k when party k is in power, the state is i, the preference shock is m_k and j is the chosen, and let $\tilde{U}_{i,\,j}^k(m_{\sim k})$ denote the current period reward to party k when party $\sim k$ is in power, the state is i, the preference shock is $m_{\sim k}$ and j is the chosen.

Assumption 2. *For all feasible $\{i, j, k\}$ triples, $u_{i,j}^k(m_k) : M \to \mathbb{R}$ and $U_{i,j}^k(m_k) : M \to \mathbb{R}$ are continuous in m_k and $\tilde{u}_{i,j}^k(m_{\sim k}) : M \to \mathbb{R}$ and $\tilde{U}_{i,j}^k(m_{\sim k}) : M \to \mathbb{R}$ are continuous in $m_{\sim k}$.*

Assumption 3. *Let $\{i, j, k\}$ and $\{i, j', k\}$ be any pairs of feasible triples. Then $U_{i\,j}^k(m_k) - U_{i\,j'}^k(m_k)$ is strictly monotone in $m_k \in M$.*

For the models in the main text, Assumption 3 is satisfied by virtue of the concavity of $U(x)$ and $U(\tau - x)$. To see this, consider $k = D$ and let $\delta(m_k) \equiv U_{i\,j}^D(m_k) - U_{i\,j'}^D(m_k)$. Then (recalling that m_D is denoted m_1 in the main text),

$$\delta(m_1) = U(g_j + m_1) + \alpha U(\tau - g_j) - \psi(g_i, g_j) - U(g_{j'} + m_1) - \alpha U(\tau - g_{j'}) + \psi(g_i, g_{j'}).$$

Observe that $\delta'(m_1) = U'(g_j + m_1) - U'(g_{j'} + m_1)$. Since $g_j \neq g_{j'}$, $g_j + m_1$ is either less than or greater than $g_{j'} + m_1$ for all m_1. By concavity of U, $\delta'(m_1)$ is either strictly positive for all m_1 or strictly negative for all m_1 (an analogous argument establishes the result for $k = R$).

[28] In the main text, m_D is m_1 (the preferred good of type D) and m_R is m_2 (the preferred good of type R).

Probability Spaces:

Let $(M, \mathcal{B}_M, \mu)$ denote a probability space on M, where $\mathcal{B}_M$ denotes the Borel σ-algebra on $\mathbb{R}$ restricted to M.

Assumption 4. *μ is absolutely continuous with respect to the Lebesgue measure on M.*

This assumption means that any subset of M that is of Lebesgue measure 0 has probability zero with respect to the probability measure μ. Any random variable described by a continuous density on M will satisfy this assumption.

Let $(\mathbb{R}, \mathcal{B}, \lambda)$ denote the probability space on $\mathbb{R}$ for the aggregate voter preference shock A.

Assumption 5. *For all $z \in \mathbb{R}$, $\int_{A>z} A\, d\lambda$ exists and is continuous in z and there is a $\overline{A} > 0$ such that $|\int_{A>z} A\, d\lambda| < \overline{A}$.*

This assumption is satisfied by any random variable with a continuous density on a compact support. It is also satisfied by a random variable with unbounded support if it possessed a continuous density over $\mathbb{R}$ and that density converged to 0 exponentially fast as A diverged to $+\infty$ or $-\infty$ (as is the case, for instance, with the normal distribution).

Let $(\mathbb{R}, \mathcal{B}, \varepsilon)$ denote a probability space on $\mathbb{R}$ for the idiosyncratic voter preference shock e and let $F(e) = \varepsilon(-\infty, e])$ be its distribution.

Assumption 6. *$\int_{\mathbb{R}} e d\varepsilon = 0$, $\varepsilon((-e, 0]) = \varepsilon([0, e))$ for all $e \in \mathbb{R}$ and $F(e)$ is continuous and strictly increasing in e.*

Any random variable with a density function that is symmetric around 0 and is strictly positive for all $e \in \mathbb{R}$ (such as the normal distribution) will satisfy Assumption 6.

In what follows, we use some standard results from measure theory and functional analysis. In most instances, the proofs of these results can be found in Stokey and Lucas Jr. (1989), and when this is the case, we cite the relevant section of their text (for instance SL, Ch. 7, p. 192).

B.2 Preliminary Lemma

Lemma 1 (Boundedness of current period rewards)**.** *There exists $\overline{U} > 0$ such that $|u^k_{ij}(m_k)|$, $|U^k_{ij}(m_k)|$, $|\tilde{u}^k_{ij}(m_{\sim k})|$, and $|\tilde{U}^k_{ij}(m_{\sim k})|$ are all strictly less than $\overline{U}$ for all feasible triples $\{i, j, k\}$ and all m_k, $m_{\sim k} \in M$.*

Proof. Since a real-valued continuous function on compact set must be bounded, each of the functions in the Lemma can be given a bound. And since there are finite number of such functions, an $\overline{U} > 0$ exceeding all of the individual bounds exists. □

B.3 Decision Problem of the Party in Power

To recall: A period begins in some state i. The voter preference shocks e and A are realized and people vote. If party k wins, the preference shock m_k is realized ($m_{\sim k}$ is automatically zero) and the party chooses next period's state j.

Let $Q_i^k \in \mathbb{R}$, $k \in \{D, R\}$, denote the value of party k of starting a period in state i.

Let $A_i \in \mathbb{R}$ denote the threshold value of A in state i, i.e., if $A > A_i$, D party wins the election in state i.

Let ω_i denote the 3-tuple $\left(Q_i^D, Q_i^R, A_i\right)$. Let $\omega = (\omega_1, \omega_2, \ldots, \omega_I)$ be a vector composed of these 3-tuples. Then ω is an element of $\{\mathbb{R} \times \mathbb{R} \times \mathbb{R}\}^I$.

We use $Q_i^D[\omega]$, $Q_i^R[\omega]$, and $A_i[\omega]$ to denote the specific elements of the ith component of ω.

Since m_k shocks are never active simultaneously, we reduce the notational burden by using m to denote realizations of whichever m_k shock is active.

Value Functions:

With these conventions, let $V_{i\,j}^k(m;\omega)$ be the value to party k when it is in power, the state is i, its preference shock is m, and it chooses j. Then,

$$V_{i,j}^k(m;\omega) = U_{i,j}^k(m) + \beta Q_j^k[\omega], \;\; k \in \{D, R\}. \tag{27}$$

Let $V_i^k(m;\omega)$ be the *optimal* value of party k under the same circumstances. Then,

$$V_i^k(m;\omega) = \max_{j \in \Gamma_i^k} V_{i,j}^k(m;\omega). \tag{28}$$

By Assumption 1, the set of maximizers is nonempty for all i. Let $j_i^k(m;\omega)$ denote the maximizer if it is unique. If the set of maximizers is not unique, we adopt the following tie-breaking rule: $j_i^k(m;\omega)$ is the maximizer with the smallest (index) j.

Proposition 2 (Continuity of V_i^k)**.** *For all i and k, $V_i^k(m;\omega) : M \times \Omega \to \mathbb{R}$ is continuous in m and ω.*

Proof. $V_i^k(m;\omega)$ is the upper envelope of a finite number of functions $V_{i,j}^k(m;\omega) : M \times \Omega \to \mathbb{R}$, $j \in \Gamma_i^k$. By Assumption 2, U_{ij}^k is continuous in m and $Q_j^k[\omega]$ is trivially continuous in ω, and so each of these functions is continuous in m and ω. Then the upper envelope of these functions must also be continuous in m and ω. □

Proposition 3 (Integrability of V_i^k w.r.t. m)**.** *Given ω, $V_i^k(\omega) \equiv \mathbb{E}_m V_i^k(m;\omega)$ exists.*

Proof. Since $V_i^k(m;\,\omega)$ is continuous in m, $V_i^k(m;\,\omega)$ is measurable with respect to $\mathcal{B}_M$ [SL, Ch. 7, p. 178]. Since M is compact, $\inf_m V_i^k(m;\,\omega)$ and $\sup_m V_i^k(m;\,\omega)$ both exist. Then there is some $\cancel{V}(\omega) > 0$ for which $\mid V_i^k(m;\,\omega) \mid < \cancel{V}(\omega)$. Therefore, $V_i^k(m;\omega)$ is a bounded and measurable function and since $\mu(M)$ is finite (equal to 1), $\int V_i^k(m;\omega)d\mu = \mathbb{E}_m V_i^k(m;\omega)$ exists [SL, Ch. 7, p. 192]. □

To complete the statement of the party's decision problem, let $X_i^k(m;\omega)$ be the value to party k when it is not in power, the state is i, the preference shock of party $\sim k$ is m and party $\sim k$ chooses j optimally. Then,

$$X_i^k(m;\omega) = \tilde{U}^k_{i\, j_i^{(\sim k)}(m;\omega)}(m) + \beta Q^k_{j_i^{(\sim k)}(m;\omega)}[\omega], \;\; k \in \{D, R\}. \tag{29}$$

Observe that the value of party k when it is not in power is not the maximum of an optimization problem. It is, instead, pinned down by the actions chosen by the other party to maximize its own objective function. Thus it is no longer true that $X_i^k(m;\omega)$ is necessarily continuous in m and ω. An inconvenient consequence is that the integrability of $X_i^k(m;\omega)$ w.r.t. m (and of other functions that similarly depend on m via decision rules) cannot be established as easily as for $V_i^k(m;\omega)$ in Proposition 3. More information on the properties of decision rules $j_i^k(m,\omega)$ is needed.

Decision Rules:

The next three Lemmas establish key properties of decision rules.

Lemma 2 (Maximizers are almost always unique)**.** *Given k, i, and ω, $j_i^k(m;\omega)$ strictly dominates any other feasible choice, except, possibly, at a finite number of m values.*

Proof. Given k, i, and ω, the optimal choice at m is unique if $V^k_{i,\,j^k_i(m;\omega)}(m;\omega) > V^k_{i,\,j'}(m;\omega)$ for all $j' \in \Gamma^k_{i,\,j} \setminus j^k_i(m;\omega)$. We show that this inequality holds for all but a finite number (possibly zero) of m values.

Let $j, j' \in \Gamma^k_i$, $j \neq j'$, and let $M^k_i(j, j'; \omega) \subseteq M$ be $\{m : V^k_{i,\,j}(m;\omega) = V^k_{i,\,j'}(m;\omega)\}$. Now,

$$V^k_{i,\,j}(m;\omega) - V^k_{i,\,j'}(m;\omega) = U^k_{i,j}(m) - U^k_{i,j'}(m) + \beta Q^k_j[\omega] - \beta Q^k_{j'}[\omega].$$

By Assumption 3, the r.h.s. is strictly monotone in m. Therefore, either $M^k_i(j, j'; \omega)$ is empty or it contains exactly one point.

Given ω, let

$$M^k(\omega) = \left\{ \bigcup_{i \in \mathcal{I}} \left\{ \bigcup_{j,\,j' \in \Gamma^k_i,\, j \neq j'} M^k_i(j, j'; \omega) \right\} \right\}$$

be the collection of all such (indifference) points for party k. Since $\mathcal{I}$ is a finite set, $M(\omega)$ is a finite set. Now consider $\hat{m} \in M \setminus M^k(\omega)$. Then $V^k_{i,\,j^k_i(\hat{m};\omega)}(m;\omega) > V^k_{i,\,j'}(\hat{m};\omega)$ for any $j' \in \Gamma^k_{i,j} \setminus j^k_i(m;\omega)$. If not, $\hat{m}$ must belong to $M^k_i(j^k_i(\hat{m};\omega), j'; \omega)$ for some j' and so must belong to $M^k(\omega)$, which is impossible in view of the choice of $\hat{m}$. Since $M \setminus M(\omega)$ contains all but a finite number of m values, the result follows. $\square$

Lemma 3 (Measurability w.r.t. m)**.** *Given i, k, and ω, let $B^k_{i,\,j}(\omega) \subseteq M$ be the set $\{m \in M : j^k_{i\,j}(m;\omega) = j\}$ of m values for which the optimal choice of party k is j. Then, $B^k_{i\,j}(\omega) \in \mathcal{B}_M$ for all $j \in \mathcal{I}$.*

Proof. We prove this by showing that $B^k_{i\,j}(\omega)$ is the union of two Borel sets.

Fix k, i, and ω. For each $j \in \mathcal{I}$, let

$$V^k_{i \setminus j}(m;\omega) = \max_{j' \in \mathcal{I} \setminus j} V^k_{i,\,j'}(m;\omega)$$

denote the optimal value of party k excluding policy j. Now consider the difference function $f^k_{i\,j}(m;\omega) : M \to \mathbb{R}$ defined as $V^k_{i\,j}(m;\omega) - V^k_{i \setminus j}(m;\omega)$. Then $\hat{B}^k_{i\,j}(\omega) = \{m \in M : f^k_{i\,j}(m;\omega) > 0\}$ is the set of m points for which j is the unique maximizer. Since $f^k_{i\,j}$ is the difference of two functions continuous in m, $f^k_{i\,j}$ is continuous in m and, hence, $\hat{B}^k_{i\,j}(\omega) \in \mathcal{B}_M$.

Next, given k, i, and ω, consider the set of m values for which the maximizer is not unique. By Lemma 2, this set is finite. Of this finite set of m values, let $\Phi^k_{i,j}(\omega)$ be the subset of m values for which $j^k_i(m;\omega) = j$, i.e., the subset of m values for which j is the optimal choice because it was the smallest j among all optimal j's (the tie-breaking rule).

Then

$$B^k_{i,j} = \hat{B}^k_{i,j} \cup \Phi^k_{i,j}(\omega).$$

Since any finite subset of M is a Borel set and the union of two Borel sets is Borel, it follows that $B^k_{i\,j}(\omega) \in \mathcal{B}_M$. □

Lemma 4. *Given i, k, and ω, let $\chi_{B^k_{i\,j}(\omega)}(m)$ denote the indicator function that is 1 if $m \in B^k_{i\,j}(\omega)$ and 0 otherwise. Let $\theta(m) : M \to \mathbb{R}$ be a continuous real-valued function of m. Then the product function $\theta(m)\chi_{B^k_{i\,j}(\omega)}(m)$ is measurable with respect to $\mathcal{B}_M$ and integrable with respect to μ.*

Proof. Since $B^k_{i\,j}(\omega) \in \mathcal{B}_M$ (Lemma 3), $\chi_{B^k_{i\,j}(\omega)}(m)$ is a measurable function. Since $\theta_j(m)$ is continuous, it is also a measurable function. Therefore, $\theta_j(m)\chi_{B^k_{i\,j}(\omega)}(m)$ (being the product of measurable functions) is also a measurable function. Since M is compact, the function $\theta(j)$ is bounded and, therefore, so is $\theta_j(m)\chi_{B^k_{i\,j}(\omega)}(m)$. Since bounded measurable functions are integrable, $\theta_j(m)\chi_{B^k_{i\,j}(\omega)}(m)$ is integrable. □

Lemma 5 (Almost everywhere convergence of decision rules)**.** *Let $\{\omega_n\}$ be a sequence converging to ω. Then, for each i and k, the sequence of functions $\{j^k_i(m;\omega_n) : M \to \mathcal{I}\}$ converges pointwise to the function $j^k_i(m;\omega) : M \to \mathcal{I}$ except, possibly, for a finite number of m values.*

Proof. Pick a point in $\hat{m} \in M$ and suppose that $j^k_i(\hat{m};\omega)$ is a unique maximizer. Let $V^{k-}_i(\hat{m};\omega) = \max_{j \in \Gamma^k_i \setminus j^k_i(\hat{m};\omega)} V^k_{i\,j}(\hat{m};\omega)$. Then, $V^k_i(\hat{m},\omega) > V^{k-}_i(\hat{m};\omega)$. Since both $V^{k-}_i(\hat{m};\omega)$ and $V^k_i(\hat{m};\omega)$ are continuous in ω, there exists N such that for all $n > N$, $V^k_i(\hat{m};\omega_n) > V^{k-}_i(\hat{m};\omega_n)$. Then $j^k_i(\hat{m};\omega_n) = j^k_i(\hat{m};\omega)$ for all $n > N$. But this implies that $\lim_n j^k_i(\hat{m},\omega_n) = j^k_i(\hat{m},\omega)$. Since the maximizer $j^k_i(m;\omega)$ is unique for all but a finite number of m's (Lemma 2), the result follows. □

Lemma 6 (Continuity of expected value)**.** *Let $\theta(m) : M \to \mathbb{R}$ be a continuous real-valued function of m. Then, $\int \theta(m)\chi_{B^k_{i\,j}(\omega)}(m)d\mu$ is continuous in ω.*

Proof. Let $\omega_n \to \omega$

$$\theta(m)\chi_{B^D_{i,j}(\omega_n)}(m) = \begin{cases} \theta(m) & \text{if} j = j^D_i(m;\omega_n) \\ 0 & \text{if} j \neq j^D_i(m;\omega_n). \end{cases}$$

By Lemma 5, $j^D_i(m,\omega_n)$ converges pointwise to $j^D_i(m,\omega)$ except, possibly, for a finite number of m values. Therefore, $f_n(m) \equiv \theta(m)\chi_{B^D_{i,j}(\omega_n)}(m)$ converges pointwise to $f(m) \equiv \theta(m)\chi_{B^D_{i,j}(\omega)}(m)$ except, possibly, for a finite number of m values. A finite subset of M has Lebesgue measure 0 and so by Assumption 4 $f_n(m)$ converges to $f(m)$, μ - a.e. Furthermore, $|f_n(m)| \leq |\theta(m)|$ is a sequence of bounded functions. By the Lebesgue Dominated Convergence Theorem (SL, Theorem 7.10, p. 192), $\lim_n \int \theta(m)\chi_{B^D_{i,j}}(\omega_n)(m)d\mu = \int \theta(m)\chi_{B^D_{i,j}}(\omega)d\mu$. □

We close this subsection with:

Proposition 4 (Integrability of X^k_i w.r.t. m)**.** *For each k, i, and ω, $X^k_i(\omega) \equiv \mathbb{E}_m X^k_i(m;\omega)$ exists.*

Proof. The key step is simply the observation that $X^k_i(m;\omega)$ can be expressed as:

$$X^k_i(m;\omega) = \sum_j \left\{\tilde{U}^k_{ij}(m) + \beta Q^k_{ij}[\omega]\right\} \chi_{B^{\sim k}_{ij}(\omega)}(m). \tag{30}$$

For each j, $\tilde{U}^k_{ij}(m)+\beta Q^k_{ij}[\omega]$ is a continuous function of m (Assumption 2) and, therefore, $\{\tilde{U}^k_{ij}(m)+\beta Q^k_{ij}[\omega]\}\chi_{B^{\sim k}_{ij}(\omega)}(m)$ is a integrable function of m (Lemma 4). Since a finite sum of integrable functions is also integrable, $X^k_i(m;\omega)$ is integrable. □

B.4 Lifetime Utilities of Voters

The goal of this subsection is to establish that given a pair of decision rules $j^k_i(m;\omega)$, $k \in \{D,R\}$, and values for the thresholds $A_i[\omega]$ of the aggregate voter preference shock is (above which the D party is elected), the voter value functions $\{W^k_i(m;\omega), Z^k_i(m;\omega)\}$ $k \in \{D,R\}$ are uniquely determined.[29] Furthermore, these value functions are continuous in ω.

Let $\Pi_i(\omega) \equiv \int_{z>A_i[\omega]} d\lambda$ denote the probability of a D party win given ω and let $\overline{A}_i(\omega) \equiv \int_{z>A_i[\omega]} A\, d\lambda$ denote the A_i-truncated-expectation of A, which exists by Assumption 5.

[29]Recall that these value functions give the lifetime utility of type k when the state is i and (the active) preference shock is m, ignoring the value of an individual's net preference for the D party when the D party is in power).

Then, the voter value functions, if they exist, must satisfy the following recursions:

$$W_i^D(m;\omega) = \tag{31}$$
$$u^D_{i,\,j_i^D(m;\omega)}(m) + \beta\left\{\Pi_{j_i^D(m;\omega)}\mathbb{E}_{m'}W^D_{j_i^D(m;\omega)}(m';\omega) + \left[1-\Pi_{j_i^D(m;\omega)}\right]\mathbb{E}_{m'}Z^D_{j_i^D(m;\omega)}(m';\omega) + \overline{A}_{j_i^D(m;\omega)}\right\}$$
$$Z_i^D(m;\omega) = \tag{32}$$
$$\tilde{u}^D_{i,\,j_i^R(m;\omega)}(m) + \beta\left\{\Pi_{j_i^R(m;\omega)}\mathbb{E}_{m'}W^D_{j_i^R(m;\omega)}(m';\omega) + \left[1-\Pi_{j_i^R(m;\omega)}\right]\mathbb{E}_{m'}Z^D_{j_i^R(m;\omega)}(m';\omega) + \overline{A}_{j_i^R(m;\omega)}(\omega)\right\},$$

and

$$W_i^R(m;\omega) = \tag{33}$$
$$u^R_{i,\,j_i^R(m;\omega)}(m) + \beta\left\{\Pi_{j_i^R(m;\omega)}\mathbb{E}_{m'}Z^R_{j_i^R(m;\omega)}(m';\omega) + \left[1-\Pi_{j_i^R(m;\omega)}\right]\mathbb{E}_{m'}W^R_{j_i^R(m;\omega)}(m';\omega) + \overline{A}_{j_i^R(m;\omega)}(\omega)\right\}$$
$$Z_i^R(m;\omega) = \tag{34}$$
$$\tilde{u}^R_{i,\,j_i^D(m;\omega)}(m) + \beta\left\{\Pi_{j_i^D(m;\omega)}\mathbb{E}_{m'}Z^R_{j_i^D(m;\omega)}(m';\omega) + \left[1-\Pi_{j_i^D(m;\omega)}\right]\mathbb{E}_{m'}W^R_{j_i^D(m;\omega)}(m';\omega) + \overline{A}_{j_i^D(m;\omega)}(\omega)\right\}.$$

Proposition 5. *Let $\mathcal{F}$ denote the set of all $\mathcal{B}_M$-measurable functions $f : M \to \mathbb{R}$ for which $\int f\,d\mu$ exists with respect to the probability space $(M, \mathcal{B}_M, \mu)$. Then, for every ω, there exists a set of functions $\left\{W_i^k(m;\omega), Z_i^k(m;\omega)\right\}$, $i \in \mathcal{I}$, all members of $\mathcal{F}$, that satisfy the recursions* (31)–(32) *for $k = D$ and* (33)–(34) *for $k = R$.*

Proof. We will prove the proposition for $k = D$ (the proof for $k = R$ is analogous).

Fix ω. We may view the r.h.s of (31)–(32) as an operator taking as input the set of functions $(W^D(m;\,\omega),\, W^R(m;\,\omega)) \equiv \{W_i^D(m;\,\omega),\, Z_i^D(m;\,\omega)\}_{i\in\mathcal{I}}$. We will establish that if all members of this set belong to $\mathcal{F}$, then the output functions on the l.h.s. of (31)–(32) also belong to $\mathcal{F}$.

First, observe that if the input functions are members of $\mathcal{F}$, then the expectations of these functions with respect to m' exist and the r.h.s. is well defined.

Next, observe that we can reexpress the r.h.s. of (31)–(32) as:

$$W_i^D(m;\omega) = \sum_j \left[u_{i,j}^D(m) + \beta\left\{\Pi_j(\omega)\mathbb{E}_{m'}W_j^D(m';\omega) + [1-\Pi_j(\omega)]\,\mathbb{E}_{m'}Z_j^D(m';\omega) + \overline{A}_j[\omega]\right\}\right]\chi_{B_{i,j}^D(\omega)}(m) \tag{35}$$

$$Z_i^D(m;\omega) = \sum_j \left[\tilde{u}_{i,j}^D(m) + \beta\left\{\Pi_j(\omega)\mathbb{E}_{m'}W_j^D(m';\omega) + [1-\Pi_j(\omega)]\,\mathbb{E}_{m'}Z_j^D(m';\omega) + \overline{A}_j[\omega]\right\}\right]\chi_{B_{i,j}^R(\omega)}(m). \tag{36}$$

By Lemma 4, each term in the summation is an integrable function of m and, therefore, the summation is as well. Hence, the output functions in the l.h.s. of (31)–(32) belong to $\mathcal{F}$.

Taking expectations w.r.t. m on both sides (35) – (36) yields a pair of recursions in the expectation (w.r.t. m) of the D types' value functions:

$$\mathbb{E}_m W_i^D(m,\omega) = \sum_j \left[\int u_{i,j}^D(m)\chi_{B_{i,j}^D(\omega)}(m)d\mu + \left[\beta\left\{\Pi_j(\omega)\mathbb{E}_{m'}W_j^D(m';\omega) + [1-\Pi_j(\omega)]\,\mathbb{E}_{m'}Z_j^D(m';\omega) + \overline{A}_j[\omega]\right\}\right]\mu(B_{i,j}^D(\omega))\right] \tag{37}$$

$$\mathbb{E}_m Z_i^D(m;\omega) = \sum_j \left[\int \tilde{u}_{i,j}^D(m)\chi_{B_{i,j}^R(\omega)}(m)d\mu + \left[\beta\left\{\Pi_j(\omega)\mathbb{E}_{m'}W_j^D(m';\omega) + [1-\Pi_j(\omega)]\,\mathbb{E}_{m'}Z_j^D(m';\omega) + \overline{A}_j[\omega]\right\}\right]\mu(B_{i,j}^R(\omega))\right]. \tag{38}$$

Or, more compactly,

$$\overline{W}_i^D(\omega) = \sum_j \left[\int u_{i,j}^D(m)\chi_{B_{i,j}^D(\omega)}(m)d\mu + \left[\beta\left\{\Pi_j(\omega)\overline{W}_j^D(\omega) + [1-\Pi_j(\omega)]\,\overline{Z}_j^D(\omega) + \overline{A}_j[\omega]\right\}\right]\mu(B_{i,j}^D(\omega))\right] \tag{39}$$

$$\overline{Z}_i^D(\omega) = \sum_j \left[\int \tilde{u}_{i,j}^D(m)\chi_{B_{i,j}^R(\omega)}(m)d\mu + \left[\beta\left\{\Pi_j(\omega)\overline{W}_j^D(\omega) + [1-\Pi_j(\omega)]\,\overline{Z}_j^D(\omega) + \overline{A}_j[\omega]\right\}\right]\mu(B_{i,j}^R(\omega))\right]. \tag{40}$$

We may verify that the operator defined by the r.h.s. of (39)– (40):

$$\left\{W_i^D(\omega)\left(\overline{W}^D(\omega), \overline{Z}^D(\omega)\right),\, Z_i^D(\omega)\left(\overline{W}^D(\omega), \overline{Z}^D(\omega)\right)\right\}_{i\in\mathcal{I}} : \{\mathbb{R}\times\mathbb{R}\}^I \to \{\mathbb{R}\times\mathbb{R}\}^I$$

satisfies Blackwell's sufficiency conditions for a contraction map (with modulus of contraction β) (SL, Theorem 3.3, p. 54). Since $\{\mathbb{R}\times\mathbb{R}\}^I$ is a complete metric space (with, say, the uniform metric), the Contraction Mapping Theorem (SL Theorem 3.2, p. 50) ensures the existence of a unique pair of vectors $(\overline{W}^{*D},\overline{Z}^{*D})$ satisfying

$$\left(\overline{W}^{*D},\overline{Z}^{*D}\right)=\left(W_i^D(\overline{W}^{*D},\overline{Z}^{*D}),Z_i^D(\overline{W}^{*D},\overline{Z}^{*D})\right)_{i\in\mathcal{I}}.$$

Then the functions, all members of $\mathcal{F}$, whose existence is asserted by the Proposition are given by:

$$W_i^D(m;\omega)=u^D_{i,j_i^D(m;\omega)}(m)+\beta\left\{\Pi_{j_i^D(m;\omega)}\overline{W}^{*D}_{j_i^D(m',\omega)}(\omega)+\left[1-\Pi_{j_i^D(m;\omega)}\right]\overline{Z}^{*D}_{j_i^D(m',\omega)}(\omega)+\overline{A}_{j_i^D(m;\omega)}\right\}$$
$$Z_i^D(m;\omega)=\tilde{u}^D_{i,j_i^R(m;\omega)}(m)+\beta\left\{\Pi_{j_i^R(m;\omega)}\overline{W}^{*D}_{j_i^R(m',\omega)}(\omega)+\left[1-\Pi_{j_i^R(m;\omega)}\right]\overline{Z}^{*D}_{j_i^R(m',\omega)}(\omega)+\overline{A}_{j_i^R(m;\omega)}(\omega)\right\}.$$

□

Proposition 6 (Continuity of $\overline{W}^{*k}$ and $\overline{Z}^{*k}$ with respect to ω)**.** *The fixed points* $(\overline{W}^{*k},\overline{Z}^{*k})$, $k\in\{D,R\}$, *vary continuously with* ω.

Proof. We will prove this for $(\overline{W}^{*D},\overline{Z}^{*D})$ (the proof for $k=R$ is entirely analogous). Since the operator $\left(W_i^D(\cdot\,,\cdot),\,Z_i^D(\cdot\,,\cdot)\right)_{i\in\mathcal{I}}$ is a contraction, it is sufficient to show that it is continuous in ω (see, for instance, Theorem 4.3.6 in Hutson and Pym (1980)). That is, *given* the vectors $(\overline{W}^D,\overline{Z}^D)$, the image $\left(W_i^D(\overline{W}^D,\overline{Z}^D),\,Z_i^D(\overline{W}^D,\overline{Z}^D)\right)$ varies continuously with ω for any i. We will show this for $W_i^D(\overline{W}^D,\overline{Z}^D)$ (the proof is analogous for $Z_i^D(\overline{W}^D,\overline{Z}^D)$).

From inspection of the r.h.s. of (37), the image will vary continuously with ω if, for each j, (i) $\overline{A}_j[\omega]$, (ii) $\int u^D_{i,j}(m)\chi_{B^D_{i,j}(\omega)}(m)d\mu$ and (iii) $\mu(B^D_{ij}(\omega))$ are continuous in ω.

Let $\omega_n\to\omega$. (i) Since $\overline{A}_j=\int_{A>A_j}Ad\lambda$ and $A_j[\omega_n]$ is just a sequence A_n converging to some A, continuity of $\overline{A}_j[\omega]$ with respect to ω is part of Assumption 5. (ii) Since $u^D_{i,j}(m)$ is a continuous function of m (Assumption 2), the result follows by setting $\theta(m)$ to $u^D_{i,j}(m)$ in Lemma 6. (iii) Since $\mu(B^D_{i,j}(\omega))=\int\chi_{B^D_{i,j}(\omega)}(m)$, the result follows by setting $\theta(m)=1$ in Lemma 6. □

B.5 Existence of a Fixed Point of the MPE Self-Map

Let

$$\begin{aligned} Q_i^D(\omega) &\equiv \left[\Pi_i(\omega)V_i^D(\omega) + [1-\Pi_i(\omega)]X_i^D(\omega)\right] \\ Q_i^R(\omega) &\equiv \left[\Pi_i(\omega)X_i^R(\omega) + [1-\Pi_i(\omega)]V_i^R(\omega)\right] \\ A_i(\omega) &\equiv \frac{[\overline{W}_i^{*R}(\omega) - \overline{Z}_i^{*R}(\omega)] - [\overline{W}_i^{*D}(\omega) - \overline{Z}_i^{*D}(\omega)]}{2}. \end{aligned}$$

Define $T(\omega) : \{\mathbb{R}\times\mathbb{R}\times\mathbb{R}\}^I \to \{\mathbb{R}\times\mathbb{R}\times\mathbb{R}\}^I$ as:

$$T(\omega) \equiv \left(Q_i^D(\omega), Q_i^R(\omega), A_i(\omega)\right)_{i\in\mathcal{I}}. \tag{41}$$

Then, given an ω^* such that $T(\omega^*) = \omega^*$, functions satisfying all the requirements of a MPE stated in Definition 1 can be constructed. An important step in this construction is the construction of the decision rules of the parties, given any ω. Even though the choice set is finite and discrete, the step is not trivial because choices have to be determined for *all* values of the continuous preference (m) shock. This construction is described in Appendix C.

To proceed, we first show that there is a compact subset $\Omega \subseteq \{\mathbb{R}\times\mathbb{R}\times\mathbb{R}\}^I$ such that $T(\Omega) \subseteq \Omega$, establish that $T(\omega) : \Omega \to \Omega$ is a continuous map, and then invoke Brouwer's Fixed Point Theorem (FPT) to assert the existence of ω^*.

To establish the existence of Ω, suppose that $Q_i^k \in [-\overline{U}/(1-\beta), \overline{U}/(1-\beta)]$ for all i, k. Then,

$$\mid V_{ij}^k(m;\omega) \mid \;\leq\; \mid U_{ij}^k(m) \mid + \beta \mid Q_j^k \mid < \overline{U} + \beta\overline{U}/(1-\beta) \;=\; \overline{U}/(1-\beta).$$

Therefore, $V_i^k(m;\omega) \in [-\overline{U}/(1-\beta), \overline{U}/(1-\beta)]$ and so $\mathbb{E}_m V_i^k(m) = V_i^k \in [-\overline{U}/(1-\beta), \overline{U}/(1-\beta)]$. The same line of reasoning shows $X_i^k \in [-\overline{U}/(1-\beta), \overline{U}/(1-\beta)]$. Thus, if $Q_i^k[\omega] \in [-\overline{U}/(1-\beta), \overline{U}/(1-\beta)]$, then $Q_i^k[T(\omega)] \in [-\overline{U}/(1-\beta), \overline{U}/(1-\beta)]$.

To establish a bound for $A_i[T(\omega)]$, we first show that $\overline{W}^{*k}$ and $\overline{Z}^{*k}$ are each contained in $[-(\overline{U}+\overline{A})/(1-\beta), (\overline{U}+\overline{A}/(1-\beta)]^I$. Observe that if $\overline{W}^k$ and $\overline{Z}^k$ belong in $[-(\overline{U}+\overline{A})/(1-\beta), (\overline{U}+\overline{A})/(1-\beta)]^I$, then

$$\mid W_i^k(\overline{W}^k, \overline{Z}^k) \mid < \overline{U} + \beta[(\overline{U}+\overline{A})/(1-\beta) + \overline{A}] < \overline{U} + \overline{A} + \beta[(\overline{U}+\overline{A})/(1-\beta)] = (\overline{U}+\overline{A})/(1-\beta),$$

and, analogously, $\mid Z_i^k(\overline{W}^k, \overline{Z}^k) \mid < (\overline{U} + \overline{A})/(1-\beta)$. Since the map $\left(W_i^k(\overline{W}^k, \overline{Z}^k), Z_i^k(\overline{W}^k, \overline{Z}^k)\right)_{i\in\mathcal{I}}$ is a contraction, the fixed points $\overline{W}^{*k}$ and $\overline{Z}^{*k}$ must each lie in $[-(\overline{U}+\overline{A})/(1-\beta), (\overline{u}+\overline{A})/(1-\beta)]^I$. Given these bounds, we may verify that $A_i[T(\omega)] \in [-(\overline{U} + \overline{A})/(1-\beta), (\overline{U} + \overline{A})/(1-\beta)]$ for all i. Since this bound holds for any $(A_i[\omega])_{i\in\mathcal{I}}$, we have, in particular, that if $(A_i[\omega])_{i\in\mathcal{I}} \in [-(\overline{U} + \overline{A})/(1-\beta), (\overline{U} + \overline{A})/(1-\beta)]^I$, then $(A_i[T(\omega)])_{i\in\mathcal{I}} \in [-(\overline{U} + \overline{A})/(1-\beta), (\overline{U} + \overline{A})/(1-\beta)]^I$.

Thus, we may take Ω to be the hypercube $[-\overline{\omega}, \overline{\omega}]^{3I}$, where $\overline{\omega} = [\overline{U} + \overline{A}]/(1-\beta)$.

To establish that $T(\omega)$ is continuous in $\omega \in \Omega$, we need only show that $Q_i^k[T(\omega)]$ is continuous in ω for each i and k, since by Proposition 6 we already know that $\overline{W}^{*k}(\omega)$ and $\overline{Z}^{*k}(\omega)$ are continuous in ω and, hence, $A_i[T(\omega)]$ is continuous in ω.

To proceed, observe that

$$\begin{aligned} Q_i^D[T(\omega)] &\equiv \left[\Pi_i(\omega)V_i^D(\omega) + [1-\Pi_i(\omega)]X_i^D(\omega)\right] \\ Q_i^R[T(\omega)] &\equiv \left[\Pi_i(\omega)X_i^R(\omega) + [1-\Pi_i(\omega)]V_i^R(\omega)\right]. \end{aligned}$$

We need to establish that $V_i^k(\omega) = \int V_i^k(m;\omega)\, d\mu$ is continuous in ω. Let ω_n be a sequence in Ω converging to $\omega \in \Omega$. By Proposition 2, $V_i^k(m;\omega_n)$ converges to $V_i^k(m;\omega)$ pointwise for all $m \in M$. Since $\mid V_i^k(m;\omega_n) \mid < \overline{U}/(1-\beta)$, by the Lebesgue Dominated Convergence Theorem the $\lim_n \int V_i^k(m;\omega_n)\, d\mu = \int V(m,\omega)\, d\mu$. Hence $V_i^k(\omega)$ is continuous in ω. An analogous argument establishes that $X_i^k(\omega)$ is continuous in ω. Finally, the continuity of $\Pi_i(\omega)$ follows from the continuity of $A_i[T(\omega)]$. Hence $Q_i^k[T(\omega)]$ is continuous in ω.

Since $T(\omega) : \Omega \to \Omega$ is continuous and Ω is compact, by Brouwer's FPT, there exists ω^* such that $T(\omega^*) = \omega^*$ and the existence of at least one pure strategy MPE is assured.

C Computation of Decision Rules

In this section, we describe how, given k, i, and ω, we compute the function $j_i^k(m;\omega) : M \to \mathcal{I}$. The following definition of *weakly preferred sets* is useful.

Definition 2 (Weakly preferred sets)**.** *Given k, i, m and ω, $P_{ij}^k(m;\omega) \subset \mathcal{I}$ is the weakly preferred set of j at m if and only if $j' \in P_{ij}^k(m;\omega)$ implies $V_{ij'}^k(m;\omega) \geq V_{ij}^k(m;\omega)$.*

The following lemma plays an important role in speeding up the computation.

Lemma 7 (Dominated choices)**.** *Let* $\underline{m} \leq m_1 < m_2 \leq \overline{m}$. *Let* $j^* = j_i^k(m_1;\omega)$. *Then any* $j \in \mathcal{I} \setminus P_{ij^*}^k(m_2;\omega)$ *is weakly dominated by* j^* *for all* $m \in [m_1, m_2]$.

Proof. Suppose there is some $m \in (m_1, m_2)$ for which there is an action $j_0 \in \mathcal{I} \setminus P_{i\,j^*}^k(m_2;\omega)$ such that $V_{i\,j_0}^k(m;\omega) > V_{i\,j*}^k(m;\omega)$. First, notice that j^* is always a member of $P_{ij^*}^k(m_2;\omega)$. Since $V_{i\,j^*}^k(m_1;\omega) \geq V_{i\,j_0}^k(m_1;\omega)$ (definition of j^*) and $V_{i\,j^*}^k(m_2;\omega) > V_{i\,j_0}^k(m_2;\omega)$ (by definition of j_0), it follows that there must be $\hat{m} \in [m_1, m)$ and another $\tilde{m} \in (m, m_2)$ for which $V_{i\,j_0}^k(\hat{m};\omega) = V_{i\,j^*}^k(m;\omega)$ and $V_{i\,j_0}^k(\tilde{m};\omega) = V_{i\,j^*}^k(\tilde{m};\omega)$. But this contradicts Assumption 3. Hence, $V_{i\,j^*}^k(m;\omega) \geq V_{i\,j_0}^k(m;\omega)$ for all $m \in [m_1, m_2]$, with the equality holding, possibly, only at m_1. $\square$

The algorithm proceeds as follows. To begin, separately sort $V_{i\,j}^k(\underline{m};\omega)$ and $V_{i\,j}^k(\overline{m};\omega)$ with respect to j in descending order. Let $\underline{j}^*$ be the highest-ranked action in the first list and $\overline{j}^*$ be the highest-ranked action in the second list. Set $j_i^k(\underline{m};\omega) = \underline{j}^*$.

The Initial Step

Case 1: $\underline{j}^* = \overline{j}^* = j^*$. Then, by Lemma 7, j^* strictly dominates all other actions for all $m \in (\underline{m}, \overline{m}]$. Hence, $j_i^k(m;\omega) = j^*$ for all $m \in (\underline{m}, \overline{m}]$ and we are done.

Case 2: $\underline{j}^* \neq \overline{j}^*$ and $P_{i\,\underline{j}^*}^k(\overline{m};\omega)$ (the weakly preferred set of $\underline{j}^*$ at $\overline{m}$) contains only two elements. Then, use bisection to determine the unique $m_1 \in (\underline{m}, \overline{m})$ for which $V_{i\,\underline{j}^*}^k(m_1;\omega) - V_{i\,\overline{j}^*}^k(m_1;\omega) = 0$ and set

$$j_i^k(m;\omega) = \begin{cases} \underline{j}^* & \text{for } m \in (\underline{m}, m_1) \\ \min\{\underline{j}^*, \overline{j}^*\} & \text{for } m = m_1 \\ \overline{j}^* & \text{for } m \in (m_1, \overline{m}], \end{cases}$$

and we are done.

Case 3: $\underline{j}^* \neq \overline{j}^*$ and $P_{i\,\underline{j}^*}^k(\overline{m};\omega)$ contains $n \geq 3$ elements, denoted $\{\overline{j}^*, j_2, \ldots, j_{n-1}, \underline{j}^*\}$. Then, use bisection to determine the indifference points $\{m_{\overline{j}^*}, m_2, m_3, \ldots, m_{n-1}\}$ at which $V_{i\,\underline{j}^*}^k(m_s;\omega) = V_{i\,j_s}^k(m_s;\omega)$, $s \in \{\overline{j}^*, 2, 3, \ldots, n-1\}$. Let $\tilde{m}$ be the minimum of this set of indifference points and let $\tilde{j}$ be the corresponding action. Then,

$$j_i^k(m;\omega) = \begin{cases} \underline{j}* & \text{for } m \in (\underline{m}, \tilde{m}) \\ \min\{\underline{j}^*, \tilde{j}\} & \text{for } m = \tilde{m} \\ \in P^k_{i\underline{j}^*}(\overline{m};\omega) \setminus \underline{j}^* & \text{for } m \in (\tilde{m}, \overline{m}] \end{cases}.$$

The top branch follows because $\underline{j}^*$ is the best choice at $\underline{m}$ and $\tilde{m}$ is the *first* m for which some other choice, namely $\tilde{j}$, gives the same utility as $\underline{j}^*$; the middle branch follows from our tie-breaking convention; and the bottom branch follows because by Lemma 7, $\tilde{j}$ dominates $\underline{j}^*$ for all $m > \tilde{m}$.

The Recursive Step

If the algorithm reaches Case 3, it returns to the Initial Step with $\underline{m} = \tilde{m}$ and $\underline{j}^* = \tilde{j}$. Note that it is legitimate to treat $\tilde{j}$ as a best choice at $\tilde{m}$ because $\tilde{j}$ gives the same utility as $\underline{j}^*$ at $\tilde{m}$ and $\underline{j}^*$ strictly dominates every other choice at $\tilde{m}$ (recall, again, that $\tilde{m}$ is the *first* m for which an action indifferent to $\underline{j}^*$ is encountered). Each return to the Initial Step adds a new m segment of the decision rule. Also, with each return to the Initial Step, there is at least one less action to evaluate (for instance, $P^k_{i\tilde{j}}(\overline{m};\omega)$ does not contain $\underline{j}^*$), so the algorithm is guaranteed to deliver the full decision rule in a finite number of steps.

Some remarks about the algorithm. First, for each k, i, and ω, the algorithm requires two initial sorts of $V_i^k(m;\omega)$ — one for $m = \underline{m}$ and one for $m = \overline{m}$. For each subsequent return to the Initial Step, no further sorting is necessary because we know that $\tilde{j}$ is a best action at $\tilde{m}$ and, since $V^k_{i\overline{j}^*}(\overline{m};\omega)$ is already sorted, we merely need to locate the position of $\tilde{j}$ in the sorted vector to determine $P^k_{i\tilde{j}}(\overline{m};\omega)$.

Second, if $P^k_{i\underline{j}^*}(\underline{m};\omega)$ has n elements, the maximum number of thresholds calculated is $(n - 1 + n - 2 + \ldots + 1) = (n^2 - n)/2$. This is a maximum because a return to the Initial Step could eliminate more than one choice. For instance, an action that is in $P^k_{i\underline{j}^*} \setminus \{\underline{j}^*, \tilde{j}\}$ may not be in $P^k_{i\tilde{j}}$. In any case, the number of thresholds calculated grows polynomially in n.[30]

Third, it is possible to speed up the algorithm by utilizing a property of the model that, while somewhat special, may hold in other applications as well. The property is that V^k_{ij} can be expressed

[30]The size n depends positively on the number of discrete choices available at each k, i, and ω (generally, this depends on the grid size of the state space) and negatively on the width of the support of m (a narrow support means that $\underline{m} \approx \overline{m}$ and so the ranking of js for $V^k_{ij}(\underline{m};\omega)$ will be quite similar to the ranking for $V^k_{ij}(\overline{m};\omega)$.

as a sum of two terms: one that depends monotonically on j and m only and a second term that depends on i and j but is independent of m. Specifically,

$$V_{ij}^{D}(m;\omega) = u(g_j + m) + B_{ij}^{D}(\omega) \quad \text{where} \quad B_{ij}^{D}(\omega) = \alpha u(\tau - g_j) - \eta(g_i - g_j)^2 + \beta Q_j^{D}[\omega]$$

and

$$V_{ij}^{R}(m;\omega) = u(\tau - g_j + m) + B_{ij}^{R}(\omega) \quad \text{where} \quad B_{ij}^{R}(\omega) = \alpha u(g_j) - \eta(g_i - g_j)^2 + \beta Q_j^{R}[\omega].$$

Since $g_{i+1} > g_i$ (by assumption), the first component is strictly increasing in j for $k = D$ and strictly decreasing in j for $k = R$, regardless of any given value of m. Focusing for the moment on the $k = D$ case, this implies that for an action j to be not dominated by an action $j' > j$, $B_{ij}^{D}(\omega) > B_{ij'}^{D}(\omega)$. If this inequality is violated, then j is strictly dominated by j' for all m and so can be dropped from further consideration. Thus, by examining the ordering of $B_{ij}^{k}(\omega)$ over j, it is often possible to prune the set of choices the algorithm has to consider.

Finally, there is a property of $j_i^k(m;\omega)$ that holds for our model and which may hold in other applications as well. We do not use this property in the computation, but its existence serves as a check on the results. This is the property of monotonicity of $j_i^k(m;\omega)$ with respect to m: For $m' > m$, $j_i^D(m;\omega) \leq j_i^D(m';\omega)$ and $j_i^R(m;\omega) \geq j_i^R(m';\omega)$. The proof follows easily from the concavity of u and we omit it.

References

Acemoglu, D., M. Golosov, and A. Tsyvinski (2011): "Power Fluctuations and Political Economy," *Journal of Economic Theory*, 146, 1009–1041.

Aghion, P., and P. Bolton (1990): "Government Debt and the Risk of Default: A Political-Economic Model of the Strategic Role of Debt," in *Public Debt Management: Theory and History*, ed. by R. Dornbusch, and M. Draghi, Chap. 11, pp. 315–345. Cambridge University Press.

Aguiar, M., and M. Amador (2011): "Growth in the Shadow of Expropriation," *Quarterly Journal of Economics*, 126, 651–697.

Ales, L., P. Maziero, and P. Yared (2014): "A Theory of Political and Economic Cycles," *Journal of Economic Theory*, 153, 224–251.

Alesina, A. (1988): "Credibility and Policy Convergence in a Two-Party System with Rational Voters," *American Economic Review*, 78(4), 796–805.

Alesina, A., and H. Rosenthal (1995): *Partisan Politics, Divided Government and the Economy*. Cambridge University Press, Cambridge, UK.

Alesina, A., and G. Tabellini (1990): "A Positive Theory of Fiscal Deficits and Government Debt in a Democracy," *Review of Economic Studies*, 57, 403–414.

Ansolabehere, S., and J. M. Snyder Jr. (2002): "The Incumbency Advantage in U.S. Elections: An Analysis of State and Federal Offices, 1942-2000," *Election Law Journal*, 1(3), 315–338.

Azzimonti, M. (2011): "Barriers to Investment in Polarized Societies," *American Economic Review*, 101(5), 2182–2204.

Azzimonti, M., M. Battaglini, and S. Coate (2016): "The Costs and Benefits of Balanced Budget Rules: Lessons from a Political Economy Model of Fiscal Policy," *Journal of Public Economics*, 136, 45–61.

Barro, R. J. (1973): "The Control of Politicians: An Economic Model," *Public Choice*, 14(1), 19–42.

Bartels, L., and J. Zaller (2001): "Presidential Vote Models: A Recount," *Political Science and Politics*, 34(1), 9–20.

Battaglini, M., and S. Coate (2008): "A Dynamic Theory of Public Spending, Taxation and Debt," *American Economic Review*, 98, 201–236.

Besley, T., and S. Coate (1998): "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis," *American Economic Review*, 88(1), 139–156.

Bowen, T., Y. Chen, and H. Eraslan (2014): "Mandatory versus Discretionary Spending: The Status Quo Effect," *American Economic Review*, 104(10), 2941–2974.

Brady, D. W., M. P. Fiorina, and A. S. Wilkins (2011): "The 2010 Elections: Why Did Political Science Forecasts Go Awry?," *PS: Political Science and Politics*, 44(2), 247–250.

Calvert, R. L. (1985): "Robustness of the Multidimensional Voting Model: Candidate Motivation, Uncertainty, and Convergence," *American Journal of Political Science*, 29(1), 69–95.

Campbell, J. E. (1997): "The Presidential Pulse and the 1994 Midterm Congressional Elections," *Journal of Politics*, 59(3), 830–857.

Cao, D., and I. Werning (2018): "Saving and Dissaving with Hyperbolic Discounting," *Econometrica*, 86(3), 805–857.

Chatterjee, S., and B. Eyigungor (2012): "Maturity, Indebtedness and Default Risk," *American Economic Review*, 102(6), 2674–2699.

——— (2016): "Continuous Markov Equilibria with Quasi-Geometric Discounting," *Journal of Economic Theory*, 163(May), 467–494.

——— (2017): "Endogenous Political Turnover and Fluctuations in Sovereign Default Risk," Working Paper 17-01, Federal Reserve Bank of Philadelphia.

Dixit, A., G. Grossman, and F. Gul (2000): "Dynamics of Political Compromise," *Journal of Political Economy*, 108(3), 531–568.

Doraszelski, U., and M. Satterthwaite (2010): "Computable Markov-Perfect Industry Dynamics," *RAND Journal of Economics*, 41(2), 215–243.

Downs, A. (1957): *An Economic Theory of Democracy*. New York: Harper and Row.

Duggan, J., and T. Kalandrakis (2012): "Dynamic Legislative Policy Making," *Journal of Economic Theory*, 147, 1653–1688.

Dziuda, W., and A. Loeper (2016): "Dynamic Collective Choice with Endogenous Status Quo," *Journal of Political Economy*, 124(4), 1148–1186.

Erikson, R. S. (1971): "The Advantage of Incumbency in Congressional Elections," *Polity*, 3(3), 395–405.

Erikson, R. S. (1988): "The Puzzle of Midterm Loss," *Journal of Politics*, 50(4), 1011–1029.

Erikson, R. S., M. B. Mackuen, and J. A. Stimson (2002): *The Macro Polity*. Cambridge University Press.

Fair, R. C. (2009): "Presidential and Congressional Vote-Share Equations," *American Journal of Political Science*, 53(1), 55–72.

Feldman, M., and C. Gilles (1985): "An Expository Note on Individual Risk without Aggregate Uncertainty," *Journal of Economic Theory*, 35, 26–32.

Ferejohn, J. (1986): "Incumbent Performance and Electoral Control," *Public Choice*, 50(1), 5–25.

Gelman, A., and G. King (1990): "Estimating Incumbency Advantage Without Bias," *American Journal of Political Science*, 34(4), 1142–1164.

Harsanyi, J. C. (1973): "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points," *International Journal of Game Theory*, 2, 1–23.

Hassler, J., J. V. R. Mora, K. Storesletten, and F. Zilibotti (2003): "Survival of the Welfare State," *American Economic Review*, 93(1), 87–112.

Hutson, V., and J. S. Pym (1980): *Applications of Functional Analysis and Operator Theory*. Academic Press.

Jacobson, G. C. (2015): "It's Nothing Personal: The Decline of the Incumbency Advantage in U.S. House Elections," *Journal of Politics*, 77(3), 861–873.

Judd, K. L. (1985): "Law of Large Numbers with a Continuum of IID Random Variables," *Journal of Economic Theory*, 35, 19–25.

Kroger, G., and M. J. Lebo (2012): "Strategic Party Government and the 2010 Elections," *American Politics Research*, 40(5), 927–945.

LEWIS-BECK, M. S., AND T. W. RICE (1984): "Forecasting U.S. House Elections," *Legislative Studies Quarterly*, 9(3), 475–486.

LINDBECK, A., AND J. WEIBULL (1987): "Balanced-Budget Redistribution as the Outcome of Political Competition," *Public Choice*, 52(3), 273–297.

MATEOS-PLANAS, X. (2012): "Demographics and the Politics of Capital Taxation in a Life-Cycle Economy," *American Economic Review*, 100(1), 337–363.

MAYHEW, D. R. (2008): "Incumbency Advantage in U.S. Presidential Elections: The Historical Record," *Political Science Quarterly*, 123(2), 201–228.

MILESI-FERRETTI, G. M., AND E. SPOLARE (1994): "How Cynical Can an Incumbent Be? Strategic Policy in a Model of Government Spending," *Journal of Public Economics*, 55(1), 121–140.

PERSSON, T., AND L. SVENSSON (1989): "Why a Stubborn Conservative Would Run a Deficit: Policy with Time-Inconsistent Preferences," *Quarterly Journal of Economics*, 104, 325–345.

PIGUILLEM, F., AND A. RIBONI (2015): "Spending-Biased Legislators: Discipline Through Disagreement," *Quarterly Journal of Economics*, 130(2), 901–949.

POOLE, K. T., AND H. ROSENTHAL (1985): "A Spatial Model for Legislative Roll Call Analysis," *American Journal of Political Science*, 29(2), 357–384.

ROGOFF, K. (1990): "Equilibrium Political Budget Cycles," *American Economic Review*, 80(1), 21–36.

SCHOLL, A. (2017): "The Dynamics of Sovereign Default Risk and Political Turnover," *Journal of International Economics*, 108, 37–53.

SONG, Z., K. STORESLETTEN, AND F. ZILIBOTTI (2012): "Rotten Parents and Disciplined Children: A Politico-Economic Theory of Public Expenditure and Debt," *Econometrica*, 80(6), 2785–2803.

STOKES, D., AND G. IVERSON (1962): "On the Existence of Forces Restoring Party Competition," *American Association for Public Opinion Research*, 26(2), 159–171.

STOKEY, N. L., AND R. E. LUCAS JR. (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press.

Tufte, E. R. (1978): *Political Control of the Economy*. Princeton University Press.

Uhlig, H. (1996): "A Law of Large Numbers for Large Economies," *Economic Theory*, 8, 41–50.

Wittman, D. (1983): "Candidate Motivation: A Synthesis," *American Political Science Review*, 77(1), 142–157.