



PERUVIAN ECONOMIC ASSOCIATION

Cause and Effect in Political Polarization: A Dynamic Analysis

Steven Callander

Juan Carlos Carbajal

Working Paper No. 173, November 2020

The views expressed in this working paper are those of the author(s) and not those of the Peruvian Economic Association. The association itself takes no institutional policy positions.

Cause and Effect in Political Polarization: A Dynamic Analysis *

Steven Callander [†] Juan Carlos Carbajal [‡]

September 9, 2020

Abstract Political polarization is an important and enduring puzzle. Complicating attempts at explanation is that polarization is not a single thing. It is both a description of the current state of politics today and a dynamic path that has rippled across the political domain over multiple decades. In this paper we provide a simple model that is consistent with both the current state of polarization in the U.S. and the process that got it to where it is today. Our model provides an explanation for why polarization appears incrementally and why it was elites who polarized first and more dramatically whereas mass polarization came later and has been less pronounced. The building block for our model is voter behavior. We take an ostensibly unrelated finding about how voters form their preferences and incorporate it into a dynamic model of elections. On its own this change does not lead to polarization. Our core insight is that this change, when combined with the response of strategic candidates, creates a feedback loop that is able to replicate many features of the data. We explore the implications of the model for other aspects of politics and trace out what it predicts for the future course of polarization.

Keywords: Political Polarization, Electoral Competition, Dynamic Analysis, Behavioral Voters.

*We have benefited from the helpful comments of Avi Acharya, Dave Baron, Gabriel Carroll, Charlotte Cavaille, Dana Foarta, Jon Eguia, Gabriele Gratton, Andy Hall, Matt Jackson, Keith Krehbiel, Andrew Little, Andrea Mattozzi, Nolan McCarty, Kirill Pogorelskiy, and seminar audiences at Stanford GSB, University of Warwick, University of Sydney, LSE, the Econometric Society Summer Meetings, the Australasian Economic Theory Workshop, and the Australian Political Economy Network. Jason Lambe provided excellent research assistance. The second author gratefully acknowledges financial support from the Australian Research Council under grant DP190101718.

[†]Graduate School of Business, Stanford University, USA. Email: sjc@stanford.edu.

[‡]School of Economics, UNSW-Sydney, Australia. Email: jc.carbajal@unsw.edu.au.

1 Introduction

Political polarization is an important and enduring puzzle. A big part of the challenge to explain polarization is that it is not just a single thing. Polarization is both a description of the current state of politics today, and a dynamic path that has rippled across the political domain over multiple decades. Polarization is, as the sociologists [DiMaggio et al. \(1996\)](#) put it, “both a state and a process.”

Adding to the complexity is that polarization’s dynamic path has not been simple. Polarization has affected different groups in different ways and to different degrees. In the United States, elite polarization has proceeded monotonically since the 1970’s, accumulating to such a substantial degree that in the U.S. Congress there remains no overlap ideologically between representatives of the two major parties. In contrast, the mass public has not polarized to the same degree, and to the extent it has polarized, the process began later and has been far less pronounced ([Gentzkow, 2016](#)).

The objective of this paper is to provide a simple model that accounts for the richness of polarization. Our model seeks to provide not only an explanation for why politics is polarized today, but why took so long to get to where it is, and why different groups have polarized to different degrees, at different speeds, and at different times. Our model adopts a classic model of electoral competition and amends it with an intuitive and behaviorally-justified change to the nature of voter preference. On its own this change is innocuous and does not lead to polarization. However, when this change is interacted with strategic candidates and iterated, the impact on politics is dramatic. We show that it leads to a rich dynamic that can account for multiple moments in the data as well as other features of voting and political behavior, thereby providing an integrated explanation of polarization across time and across different groups.

The building block for our model is the behavioral finding that preferences and behavior coexist in a causal loop. The classic view of decision making is that preferences are fixed and that they determine choice. Increasingly, evidence across many domains points to causality also running the other way: that behavior also affects preferences. In politics this causal loop works through the voting decision. Indeed, political scientists have long documented that the act of voting itself changes a citizen’s preferences. After voting for a candidate, the evidence suggests, a voter updates her preferences such that she likes that candidate a little more ([Beasley and Joslyn, 2001](#)).

We incorporate this feedback loop into a model of preferences by endogenizing a voter’s ideal point. Formally, a voter updates her ideal point by moving it toward the location of the party she voted for, even if by only a small amount. We make no other changes, retaining otherwise the classic conception of *expressive* voting with abstention.¹

¹This conception of voter behavior was first promulgated by [Hotelling \(1929\)](#) and [Smithies \(1941\)](#). We briefly discuss other possible theories of voting when we present our model in [Section 2](#).

This means that *within* each election voting behavior is standard—citizens vote for the nearest candidate or otherwise abstain—and remains consistent with the large body of evidence that has accumulated on how votes are cast.

The causal loop impacts behavior only *across* periods. Even then it does not, on its own, lead to polarization. If party positions are fixed, the feedback loop leads to a congealing of voters around the parties. Voters on the flank update inward, and voters toward the center update outward. This process produces homogenized but not necessarily polarized voting blocs. Indeed, if parties are located at moderate policies, this process produces an overall moderation of the electorate.

The key to our result is the addition of strategic parties and how they react to the evolving preferences of the citizenry. The congealing of preferences is impactful not because of the homogenization of preferences *per se* but because of the impact it has on the incentives of the parties. The updating process leads to gaps opening up in the distribution of ideal points, both between voters and abstainers and between voters for the two parties. The gap at the center of the distribution is important as it is in the center where political competition plays out. A gap in the distribution implies there are no voters to gain or lose, freeing parties to polarize toward their own, more extreme, preferences without fear of losing voters. This incentive to polarize depends only on the inside margin and exists whether the overall electorate is polarizing or moderating. This dynamic exposes how the preferences of voters and elites, despite their interdependence, can evolve at different speeds and, indeed, even in different directions.

The first step begins an iterative process of polarization. The more the parties polarize the more voters update towards them, widening the gap, and allowing the parties to polarize further, creating a feedback loop. Critically, this feedback loop is incremental and complete polarization is not immediate. The process is incremental as opening of a gap fundamentally changes the nature of political competition. The opening of a gap relaxes competition for centrist voters but it does not relax the competition that parties face with voter apathy. If the parties polarize too quickly, they risk alienating their own supporters and losing them to abstention. Electoral competition in the model shifts over time from competition for the swing voter to competition for turnout, an evolution consistent with evidence from U.S. election campaigns (Panagopoulos, 2016). This implies that the speed of polarization is tightly linked to the degree voters update their preferences after voting. The smaller, more incremental is the updating, the slower and more iterative is party polarization.

That polarization in the model is progressive illustrates how the preferences of the elites and the masses co-evolve. The equilibrium matches not only the polarization of both groups, but it explains how the timing and degree of polarization that is observed in the data can emerge. In our model it is the elites who polarize first and always lead the masses, regardless of the speed at which they polarize. The masses, in contrast, may even moderate at first before reversing course and polarizing. On the surface, this gives

the appearance of polarization being an elite-driven phenomenon, but as our model demonstrates, a necessary part of the root cause is voter preferences, without which elite polarization would not occur.

The equilibrium in our model also rationalizes other features of political behavior related to polarization. The emergence endogenously of a gap in the distribution of preferences matches the often-lamented “missing middle” of the electorate, what [Abramowitz \(2010\)](#) refers to as the “disappearing center.” In drawing a clear distinction between the preferences of voters and abstainers (abstainers don’t update their preferences), we are able to show simultaneously how those engaged with politics can polarize while their fellow citizens become increasingly disenchanted with partisan politics, leading to a bi-modal distribution of political preferences. This finding matches observation and goes some way to reconciling the competing findings about voter preferences that have riven the empirical literature.²

Behavior within our model also relates to some empirical puzzles that are ostensibly unrelated to polarization. One example is negative partisanship. Although voters update toward their favored party, the party itself is polarizing. This means that moderate voters do not, at first, get any closer to the party as they effectively chase it to the extremes during the polarization process. Consequently, despite their updating, these voters do not evaluate their favored party any higher. At the same time, these voters are moving away from the opposing party—at a rapid rate as that party moves in the opposite direction—and their evaluation of that party is declining. This creates a perplexing combination of preferences as voters seemingly become only more negative about the opposition and no less favorable about their favored party, yet this is exactly the pattern of preferences that defines “negative partisanship” and that has been extensively documented empirically ([Abramowitz and Webster, 2016](#)).

The process of polarization will not stop at the present day, requiring us to look forward as well as back. To that end, we put our model to work to explore the future of polarization. Following the logic of our model through, it predicts that party elites will continue to polarize until they reach their own ideal points where they will stabilize. That does not end voter polarization, however, as over time voters will increasingly converge on the positions of their favored party. That voters in practice are currently not as polarized as elites suggests this process still has a ways to run. This future is surely not of comfort to those who lament the state of politics today as it implies a future electorate that is as polarized as elites, with an ever larger ‘missing middle’ and with partisan constituencies that approach homogeneity.³ Moreover, at the limit of the model, the state of polarization is not only an extreme but also an increasingly stable outcome.

Throughout the baseline model, we work with symmetric parties and voters. In

²We return to this controversy in detail later in the paper.

³This is consistent and a natural end point to the evidence on increasing within-party homogeneity presented in [Levendusky \(2009\)](#)

practice, polarization in the U.S. has been asymmetric, with one side (the Republicans) polarizing faster and to a greater extent than the other. Through three simple variants, we extend the model and show several ways in which this asymmetry can emerge. We focus on the feedback loop between voters and parties, showing how micro asymmetries in voter updating and preferences can reverberate through the system to produce macro asymmetries in party positioning, both in the extent of polarization as well as the speed. As a by-product, this exercise demonstrates how path dependence and momentum effects can emerge in political outcomes, two effects that have been documented in the empirical literature but until now not connected to polarization.

The power of our result is in its simplicity. With a single, empirically-grounded change to voter behavior, the model is able to rationalize the rich dynamics of the past few decades in U.S. politics and make predictions about the future. That said, politics in practice is complicated, and definitely more complicated than is our model. We make no claim that our explanation is the only force causing polarization, whether at the level of elites or the masses, and the explanatory power of our model has its limits, to be sure. Rather, we seek to illuminate a simple mechanism that organizes many key features of the data.

The focus of our paper is on the era of polarization that began in the second half of the 20th century and has run through to the present day and, presumably, into the future. The evolution of political preferences did not begin with this era, of course, and prior to the era of polarization there was an extended period of convergence. We conclude the paper by suggesting a way to connect our model to this earlier era. By adding generational turnover to our model, it can be seen how polarization can sew the seeds of its own demise and initiate an era of moderation. Exploring more fully how our theory of polarization fits into the broader, longer term landscape offers the promise of a more general understanding of the dynamics of political behavior.

Related Literature

Arguably the central notion in the formalization of politics is the conception of the policy space and the idea of an “ideal point.” Citizens, including politicians, possess ideal points that they use to evaluate the policy positions of candidates and parties. Although unremarked upon, ideal points operate effectively as reference points. They provide meaning to the position of parties, without which party platforms, and policies more generally, would be difficult to interpret. Thus, while it is perhaps perplexing to think of ideal points as changing as a result of the vote choice, this dissonance is more a matter of terminology than substance.⁴ Viewed as reference points, the assumption that ideal points evolve is in line with emerging evidence in the behavioral literature that reference

⁴The difficulty derives from the use of the prefix “ideal” for what is, in fact, better described as a transitory point of reference.

points are endogenous (Köszegi and Rabin, 2006) and that preferences themselves are the result of outside forces (Bernheim et al., 2019).⁵ Despite this connection, we nevertheless choose to retain the standard language of citizen ideal points, adding a time subscript to highlight that they are time and context dependent.

Our model contributes to the literature on electoral competition in several ways. Our main point of departure is our focus on the updating of preferences and the dynamic path of policy. To be sure, many theoretical models of political economy seek to explain the divergence of candidates and parties from the median, but we know of no model that provides a dynamic account of that movement, nor captures the simultaneous polarization of voters.

A second contribution is to the analysis of single elections, models of which have been the focus of much of the literature. An ongoing challenge in that literature is proving equilibrium existence when parties are policy motivated (Calvert, 1985). Adding abstention complicates this challenge further. Nevertheless, we show how existence problems can be overcome. Moreover, the inclusion of abstention reveals a novel type of equilibrium in which the parties do not compete directly for the median voter, a structure that matches more closely electoral competition in practice.

Although the dynamic response of candidates in our model is novel and without relation to existing models in political economy, it bears some similarity to models of switching costs in industrial organization (Klemperer, 1987). Voters in our model face no formal cost of switching their allegiance to the opposing party, rather the switching cost emerges endogenously from the preferences of the voters and the positions of the parties. This cost changes as voters update their preferences and the parties change position, and varies across voters. For voters on the flank the effect can be ambiguous as shifting inward moves them closer to both parties, whereas for centrist voters the effect is unambiguous as movement outward toward their favored party is also movement away from the opposing party. As centrist voters are the key voters driving electoral competition, the parties exploit their outward movement and increased switching cost by polarizing toward their own preferences, just as a firm in a switching-costs model exploits its attached consumers by increasing prices. Over time, the co-movement of parties and voters reinforces this effect and the endogenous switching cost for all voters increases in intensity and at an increasing rate the more elections that are held.

By now, the question of what has caused polarization has produced a large literature. This predominantly empirical literature has eliminated many explanations (such as gerrymandering in the House of Representatives), yet a clear consensus on the underlying cause (or causes) has yet to emerge (McCarty, 2019). We differ from the literature in developing a theory of the underlying mechanism that causes polarization. This enables

⁵Our specification of the endogeneity differs from the literature. Most notably, in Köszegi and Rabin (2006) reference points are (endogenously) constructed from forward-looking expectations. In contrast, the endogeneity of ideal points in our model is backward-looking and driven by behavior.

us to not only provide an explanation of polarization *per se* but also explain the speed, timing, and differences in polarization across the different levels of politics, and to produce testable predictions about other facets of voting behavior and political outcomes that can be used to verify the theory.

The preference updating rule we apply has many possible interpretations. The most natural, and the one we carry throughout, is that it is a generalized form of cognitive dissonance. In the classic formulation of cognitive dissonance, [Festinger \(1962\)](#) argues that an individual who faces a tension between their preferences and their choice will respond by updating their preferences to remove the tension. In our setting there is no tension *per se* as the citizen always votes for the nearest of the two candidates. Yet, in the same manner as does an individual in classic cognitive dissonance theory, our citizens update their preferences to make their choice seem more secure. This generalization of cognitive dissonance is consistent with evidence from psychology ([Aronson et al., 1991](#)) as well as with the application of these ideas in political science ([Beasley and Joslyn, 2001](#)) and, explicitly so, in economics ([Mullainathan and Washington, 2009](#)). The closest paper to ours is a recent contribution by [Acharya et al. \(2018\)](#) that provides a formalization of cognitive dissonance in politics, although that paper does not consider the role of strategic candidates that is central to the equilibrium dynamic in our model.⁶

2 The Model

We develop a dynamic model of electoral competition with policy motivated candidates and abstention. In each period, $t = 1, 2, \dots$, two parties D and R compete in an election. The parties simultaneously announce policy positions, d_t and r_t , that they will implement if elected. Policies are points in the classic one dimensional policy space such that $d_t, r_t \in \mathbb{R}$. The election is decided by plurality rule.

The parties are motivated both by winning office and policy outcomes—they have mixed motivations in the classic parlance. In an abuse of notation, we denote the ideal policies for the parties by D and R , respectively, where $D < 0 < R$. The benefit to each party of winning office is $\beta \geq 0$. The period utility for D is:

$$u_t^D = \begin{cases} -|d_t - D| + \beta & \text{if } D \text{ wins} \\ -|r_t - D| & \text{if } R \text{ wins.} \end{cases}$$

The intuition behind our results does not require forward-looking behavior by the parties and, for simplicity, we focus on the extreme case in which parties discount the future

⁶[Akerlof and Dickens \(1982\)](#) is the seminal introduction of cognitive dissonance into economics. See [Penn \(2017\)](#) for an interesting application of these ideas to a formal model of political values.

completely.⁷ Also for simplicity, we focus on symmetric party preferences, $D = -R$; we leave the asymmetric case for an extension (see [Section 5.4](#)).

A continuum of citizens possess ideal points distributed in \mathbb{R} . In each election, citizens either vote for one of the two parties or they abstain. We model two types of voters. The first type, *issue* voters, vote sincerely based on proximity in the policy space. The exact form of spatial voting is not important for the mechanism that drives our results. For concreteness, we adopt the expressive form of voting that originated with [Hotelling \(1929\)](#) and was extended by [Smithies \(1941\)](#) to allow for abstention. In this model, citizens evaluate parties relative to their own ideal point and vote for the closest one. If neither of the parties is sufficiently close, the citizen is alienated and abstains. This is known as abstention-due-to-alienation. Formally, a voter with ideal point v_t votes for

$$\begin{aligned} D \text{ if } |d_t - v_t| &< |r_t - v_t| \text{ and } |d_t - v_t| \leq \lambda, \\ R \text{ if } |d_t - v_t| &> |r_t - v_t| \text{ and } |r_t - v_t| \leq \lambda, \end{aligned}$$

otherwise she abstains. If indifferent between the parties she randomizes, although this tie-breaking rule will be unimportant. The constant $\lambda > 0$ is the *region of tolerance*, beyond which a citizen prefers to abstain than express a preference for either party. This voting rule is rationalized by a simple utility function: If party $J \in \{D, R\}$ has platform p , set the utility of voting for J to be $\pi(J; v_t) = \lambda - |p - v_t|$ and the utility of abstention to zero.⁸

The key novelty of the model is how voting leads to movement in a citizen's ideal point. The updating process is as follows. For a citizen with ideal point v_t who votes for a party with platform p_t , her ideal point at election $t + 1$ becomes

$$v_{t+1} = v_t + \tau(p_t - v_t), \tag{1}$$

where $0 < \tau < 1$ is the *dissonance* parameter that dictates the speed of updating. The ideal point of an abstainer does not change. We discuss the basis for this updating rule below.

The distribution of citizen ideal points, therefore, evolves from election to election as votes are cast and ideal points are updated. Initially, the distribution of citizen ideal

⁷Allowing parties to care about the future only strengthens the dynamic underlying our results as it induces the parties to converge more to the center in the first election whilst still polarizing to the same end-point, exaggerating the polarization process. We describe the intuition for this extension later in the paper.

⁸We adopt this perspective on voting as it is simple and because expressive voting accords more naturally with behavioral voters who experience cognitive dissonance. It also fits more closely with evidence from large elections than does the strategic view of voting (that is plagued by the paradox of turnout). An alternative conception of abstention within expressive voting is for it to occur when a citizen is close to indifferent between the parties—what is known as abstention-due-to-indifference. The qualitative properties of our results would be unchanged, although the technical details would vary, particularly in the first election.

points is given by a logistic distribution with mean $\mu = 0$ and scale $\alpha > 0$. Denoting this distribution by F and the corresponding density by f , we have:

$$F(v; \alpha) = \frac{1}{1 + e^{-\frac{v}{\alpha}}} \quad \text{and} \quad f(v; \alpha) = \frac{e^{-\frac{v}{\alpha}}}{\alpha(1 + e^{-\frac{v}{\alpha}})^2}.$$

The density f is strictly increasing for all $v < 0$, symmetric around its mean of 0, and thus decreasing for $v > 0$, with full support on \mathbb{R} . We work with the logistic distribution as it simplifies the analysis and allows for closed form solutions, although it is not necessary for the main qualitative properties of our results. The necessary property is log-concavity, which also holds for the normal and many other familiar distributions (Bagnoli and Bergstrom, 2005).

The behavior of issue voters is deterministic. The standard view of elections is that outcomes are more random than this behavior allows for. The classic resolution is to add some stochastic element into the electoral process. Typically, this is done by adding an idiosyncratic noise term to voter utility such that the behavior of all voters is to an equal degree random.

We adopt a different approach in which randomness varies across groups of voters. Specifically, we suppose there is a second group of voters whose behavior is random, or at least conditioned only on features of the political landscape that are uncontrollable and even unidentifiable by the parties. This dichotomy is consistent with empirical evidence that some citizens pay attention to politics and vote spatially according to policy whereas others are essentially uninformed and seemingly cast their ballots on a whim or abstain altogether (Jessee, 2009, 2010). We refer to this second set as *noise voters*, in line with the tradition in finance of noise traders.

The combination of deterministic issue voters and stochastic noise voters leads to an election outcome that, given policy positions, is itself stochastic. To avoid the distraction of excessive notation, and in the spirit of the original reduced form approach of Calvert (1985), we suppose that the quantity and behavior of noise voters is such that a party's probability of winning an election is equal to its share of issue voters in that election. This approach reduces the complexity of our analysis, both within each election and in keeping track of the evolving distribution of issue voter ideal points over time.

We define the *state of polarization* as the spread of voter ideal points at each election and document how that evolves over time. Formally, we measure polarization as the average distance of issue voters' ideal points from the mean of the distribution. Our results are not particular to this measure of spread. Throughout our main model the distribution will be symmetric in every election and our measure of polarization is equivalent to the average distance of voters' ideal points from zero.⁹

⁹Polarization applies only to issue voters as we treat noise voters as lacking any ideology. Alternatively,

The form of the updating rule in [Equation 1](#) represents a smooth generalization of cognitive dissonance theory. Consistent with more recent evidence in both psychology and political economy, we suppose that citizens motivated by issues who cast a vote update their preferences to rationalize their decisions and make their choices seem more appealing.¹⁰ Nevertheless, the degree of dissonance, and the magnitude of a voter's response, increases the more insecure a voter is in the choice they make. In contrast, classic cognitive dissonance theory imposes a sharp transition for this phenomenon, implying that it appears only when the relative appeal of the alternatives crosses over.

The specific functional form we adopt embeds several additional modeling choices. Updating is exclusively action driven. It does not depend on the identity of the party or even whether the party wins the election. (We discuss the latter possibility later in the paper.) The specification also presumes that voters update toward the location of the party when votes are cast rather than where the party might subsequently move. This is consistent with the logic of cognitive dissonance and the feedback loop between decisions and preferences. It is also appropriate given the attention of most voters is turned on during elections and off subsequently, and resonates with the scant empirical evidence on this point (see again [Beasley and Joslyn \(2001\)](#)).¹¹

Finally, to avoid corner solutions in the first election, we impose the following conditions on parameters α and β . As is standard, full convergence occurs only if the parties value the perks of office too highly. The restriction $0 \leq \beta < 2\alpha$ is sufficient to rule out this possibility. To ensure the parties do not fully polarize at the first election, we impose a lower bound on the party ideal points. We set $R = -D > \lambda^*$, where λ^* is the voter tolerance level that solves the following hyperbolic equation:¹²

$$\tanh\left(\frac{\lambda^*}{\alpha}\right) = \frac{4\alpha}{2\lambda^* + \beta}. \quad (2)$$

3 The First Election

With policy motivated parties and uncertainty over the election outcome, the first election presents the parties with a classic trade-off between the probability of winning and

we could endow them with ideologies (that they ignore in their vote choice). Assuming noise voters don't then update their ideal points, this would slow down any change in polarization.

¹⁰See the references in the Introduction.

¹¹An alternative dynamic linkage is to keep ideal points fixed and suppose that voters update the valence of the party they support. With only two parties this formulation would lead to the same dynamic polarization of elites as generated in our model but, clearly, it would not generate polarization of the masses. That said, because most surveys of spatial preferences do not include a valence dimension, the empirical connection between this formulation and polarization is unclear and in need of further research.

¹²At its tightest for the boundary value of $\beta = 0$ and setting $\alpha = 1$, this requires only that $R > 2.065$. This implies that up to 25% of the population have ideal points outside those of the parties.

the policy outcome. By inching toward the center, a party increases the chance it wins the election, but at the cost of a less attractive policy should it win. As has been known since the seminal contribution of [Calvert \(1985\)](#), this trade-off leads to an equilibrium in which the two parties do not fully converge to the center as long as the pure benefit of winning office, β , is not too large.

In the classic formulation, the competitive tension plays out exclusively at the center of the distribution with the parties competing intensely for the median voter. The logic depends, however, on full turnout. With full turnout, every citizen votes and the only competitive margin is halfway between the parties where the swing voters sit. Adding abstention changes this. If the parties' positions are far enough apart, the intervals of their support do not intersect. This leaves abstainers in the middle of the distribution, and multiple margins at which citizens are indifferent between abstaining and voting for one or other party. In this case there is, however, no margin at which citizens are indifferent between the parties and they actually turn out to vote.

This formulation has not been analyzed previously in the literature, even for one-shot elections. We show that it is important as it leads to a new type of equilibrium, one in which competition is between parties and abstention rather than the parties directly. In this equilibrium, parties stop converging before their intervals of support meet and intense competition at the center of the distribution does not occur.

[Proposition 1](#) establishes that the possibility of this new equilibrium coexists with the traditional equilibrium in which parties compete for the median voter at the center. The equilibria are distinguished by the level of voter tolerance λ . For high voter tolerance, the parties converge sufficiently such that they compete in the center, whereas for low voter tolerance centrist citizens abstain and the parties appeal to very distinct constituencies. The inclusion of abstention—and the new type of equilibrium—complicates the analysis considerably as the parties' objective functions now are only piecewise differentiable and not necessarily quasi-concave. Nevertheless, we are able to establish the uniqueness of a symmetric equilibrium for each set of parameter values and show there is a unique cut-point demarcating the two types of equilibrium.

Proposition 1. *In the first election, a unique symmetric equilibrium exists with $r_1^* = -d_1^* \in (0, R)$. The parties win election with equal probability. Party R 's equilibrium location r_1^* is implicitly defined by:*

$$1 = \left(\frac{r_1^*}{2\alpha} + \frac{\beta}{4\alpha} \right) \left\{ \tanh \left(\frac{r_1^* + \lambda}{2\alpha} \right) + \tanh \left(\frac{r_1^* - \lambda}{2\alpha} \right) \right\}, \quad \text{for } \lambda \leq \lambda^*, \quad (3)$$

$$1 = \left(\frac{r_1^*}{2\alpha} + \frac{\beta}{4\alpha} \right) \tanh \left(\frac{r_1^* + \lambda}{2\alpha} \right), \quad \text{for } \lambda > \lambda^*, \quad (4)$$

where λ^* is the tolerance level implicitly defined in [Equation 2](#).

The equilibrium policy positions of the parties are in closed-form, although only implicitly and this makes interpretation difficult. Clarity can be obtained graphically. Figure 1 depicts the two cases that are possible. For lower levels of voter tolerance, $\lambda \leq \lambda^*$, the parties stop converging before their intervals of support intersect. As a result, citizens abstain on either flank as well as in the middle. This is the solution given by Equation 3 and is depicted in the left panel of Figure 1. For larger voter tolerance, $\lambda > \lambda^*$, the intervals of support do intersect, leaving abstainers only on the flanks. This solution is given by Equation 4 and depicted in the right-side panel of the figure.

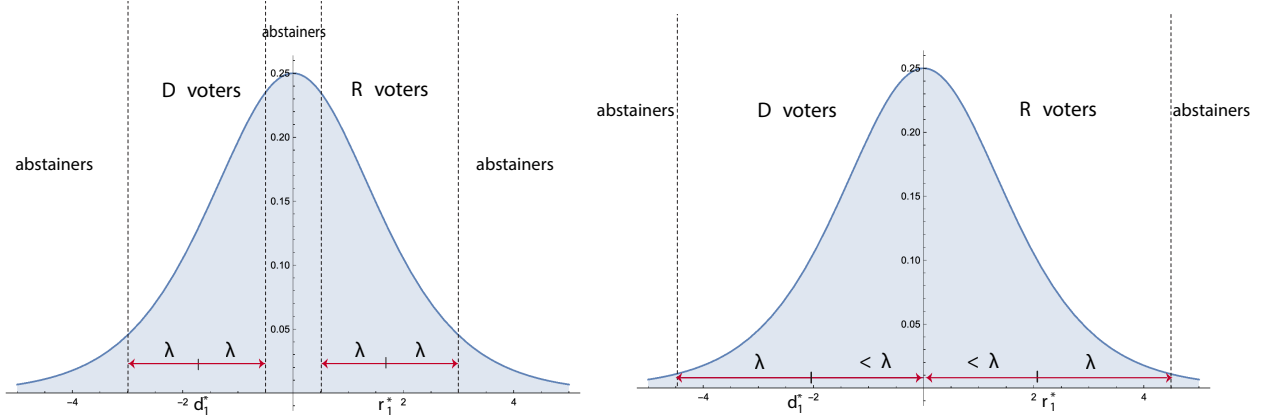


Figure 1: Equilibrium Configurations of Voters and Abstainers

The two configurations possible in equilibrium resonate with the prominent debate in political science over whether it is better for parties to appeal to their base or to swing voters (Hall and Thompson, 2018). For low voter tolerance, the parties seemingly abandon efforts to persuade voters to vote for them rather than the opposing party, concentrating exclusively on voters on the flank who would otherwise abstain. For high voter tolerance, the parties do seek to persuade as well as mobilize voters, and they compete head-to-head for centrist voters. This result shows how these two strategic options, rather than being fundamentally in contrast, can in fact emerge from a single model of electoral competition, differentiated only by parameter values. In both types of equilibrium the competitive tension is the same: creep inward for more voters at the expense of a worse policy. The novelty of the low voter tolerance equilibrium is simply that this competitive drive can exhaust itself well before the battle is met with the other party and instead resemble a mobilize-the-base strategy. We will see in later sections that this type of equilibrium, rather than being a peculiarity, in fact emerges over time as the dominant style of electoral competition, matching the dominance of the mobilize-the-base strategy over time in practice (Panagopoulos, 2016).

To better see the equilibrium, and the continuity between the two forms of competition, Figure 2 depicts the equilibrium positions as a function of voter tolerance, λ , for three different values of β , the direct benefit of winning office. The striking feature is that

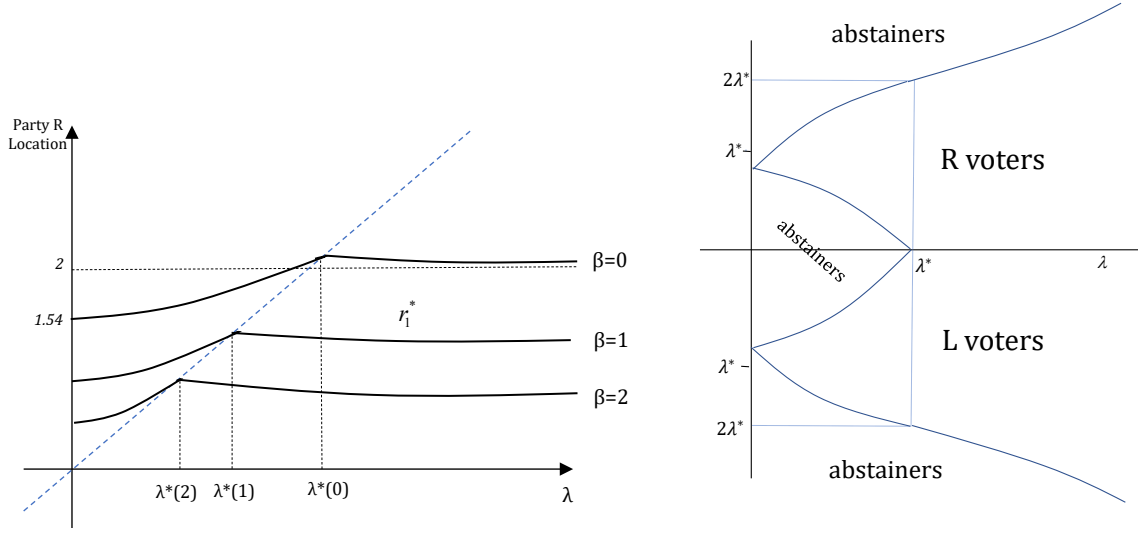


Figure 2: (i) Equilibrium Platforms in the First Election as a Function of λ . (ii) Regions of Turnout as λ Varies for $\beta = 0$

the equilibrium is not monotonic in λ . The preceding discussion suggests that higher voter tolerance leads to more direct competition between the parties, yet for low values of λ the parties actually diverge as λ increases. Nevertheless, even as they diverge, their competitive margin converges toward the median citizen. These possibilities are mutually consistent as the rate of divergence in the party positions is sufficiently slow that the inside boundary of a party's support continues to get closer to 0. At the critical threshold, λ^* , the party positions hit the 45° line, and with $\lambda = r_1^*$ the boundaries of each party's support touch at 0. This is evident in the right-side panel of Figure 2 that shows, for $\beta = 0$, how the intervals of support for the two parties grow as λ increases.

Beyond the threshold of λ^* , increases in voter tolerance have the opposite effect. With head-to-head competition now engaged between the parties, further increases in voter tolerance induce them to compete more intensely and they converge in their positions, although the degree of convergence is tempered. The limiting case as $\lambda \rightarrow \infty$ corresponds to full turnout. The non-monotonicity in party positions implies that the effect of abstention on polarization is ambiguous. Moderate and high levels of voter tolerance (but less than ∞) induce more extreme party positions than when turnout is complete. In contrast, low levels of voter tolerance lead to more moderate positions as parties seek to capture higher-density regions of citizens.

The changing rate of turnout as voter tolerance changes is depicted in Figure 3 for β values of 0 and 1. Intermediating the relationship between voter tolerance and turnout is the strategic behavior of the parties. As λ increases from low levels, growth in turnout is tempered by the increasing divergence of the party positions. A kink appears at the threshold λ^* as beyond this point new voters appear only on the flanks where density

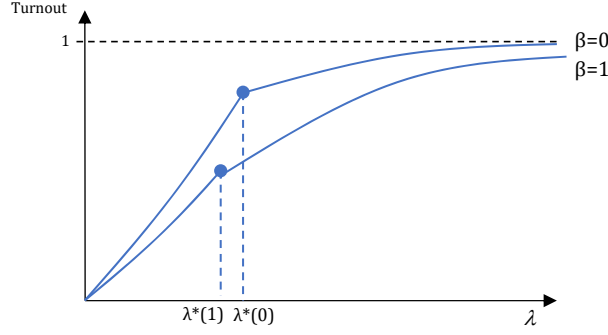


Figure 3: Turnout as λ Varies

is low, and the convergence of the parties slows this growth even further. Inevitably, however, as λ grows large, turnout grows and in the limit it approaches full turnout.

The first election equilibrium provides the starting point for the polarization that follows. It can be set to be moderate or polarized by varying the parameters α , β , and λ . An increase in β induces the parties to compete more intensely and choose more centrist positions, moving the threshold, λ^* closer to the center. The parameter α represents the scale, or variance, of the logistic distribution. As it increases, spreading the citizens' ideal points wider, the incentive of the parties to converge is muted as there are fewer voters to gain in the center. In this case, the party positions are more divergent, and so too is the critical threshold λ^* .

The formal derivation of the first-period equilibrium given in [Proposition 1](#) is complex and we relegate the details to [Appendix A](#) (including the comparative statics discussed here).¹³ As noted above, the logistic distribution simplifies our analysis but is not necessary for the main qualitative properties of equilibrium. The necessary property is log-concavity, which also holds for the normal and many other familiar distributions ([Bagnoli and Bergstrom, 2005](#)). Our key observation is that the expected utility for the parties is a product function (see [Equation 8](#) in the Appendix) and, as we show, both components of this function are log-concave. This allows us to establish the existence and uniqueness of an optimal location for each party. In this way we overcome the problem of second order conditions that normally plague models of this sort. We restrict our attention to the logistic distribution as it can be represented via hyperbolic functions. It is this feature that enables us to obtain closed form solutions.

4 Fixed Party Locations: A Benchmark

After the first election the winning party is installed in office and voters update their preferences. This changes the distribution of ideal points in two ways: the ideal points

¹³Derivations of the results for the dynamic election are collected in [Appendix B](#).

of voters compress toward the party positions and gaps open up. The gaps appear because voters update whereas abstainers don't, such that at the margin between them a discontinuity is created. If there are no abstainers in the middle then the gap is between the voters themselves as D voters shift left and R voters shift right. The two possible configurations are depicted in Figure 4. In both cases, the compression in ideal points of voters leads to higher density in those regions.

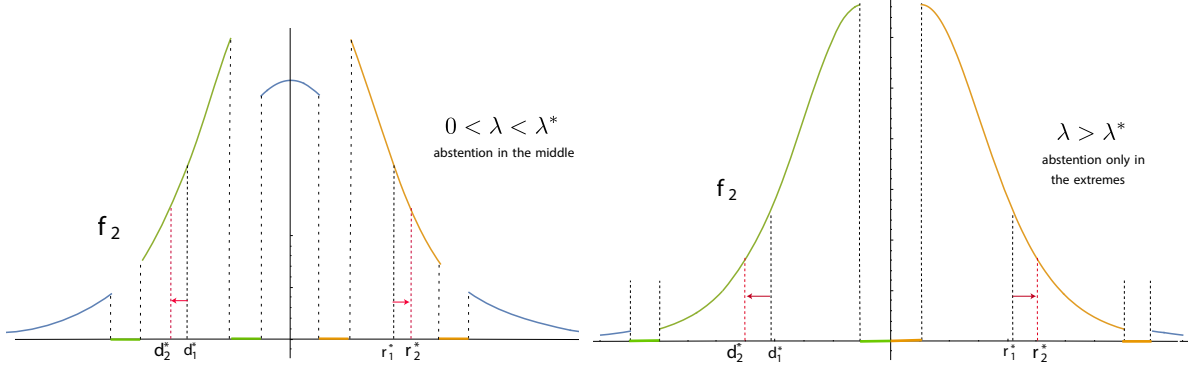


Figure 4: Ideal Point Updating for (i) Low λ , (ii) High λ

This process then iterates over time. In the full model, the evolution of voter preferences interacts with the strategic response of parties. To understand the forces at play, we begin by disentangling these effects. We fix the party positions at \hat{r} and \hat{d} for all elections and focus exclusively on voters.

With fixed party positions, the evolution of voter preferences is straightforward. Voters compress around the position of their preferred party, becoming increasingly homogeneous. At the same time, abstainers remain unmoved and are never tempted in off the sidelines to vote, leaving the turnout rate constant. The combination of these two facts implies that, over time, the gap between voters and abstainers grows ever larger. We collect these properties in the following proposition, whose proof is immediate.

Proposition 2. Fix the party positions at $\hat{r} = -\hat{d} > 0$. At each election $t \geq 2$,

- (i) the ideal point of a voter evolves monotonically, converging on \hat{r} or \hat{d} as $t \rightarrow \infty$;
- (ii) the average distance between co-partisans (who vote for the same party) decreases, approaching 0 as $t \rightarrow \infty$;
- (iii) the minimum distance between a voter and an abstainer increases from 0 and approaches λ as $t \rightarrow \infty$; and
- (iv) turnout is constant throughout elections.

The evolution of individual ideal points changes the degree of polarization at the electorate level. The aggregate effect on polarization, however, is ambiguous, depending

on the relative position of the parties and the degree of voter tolerance, λ . The key measure in determining the dynamic of polarization is the initial average ideal point of a party's voters and how that compares to the party's position itself. The critical threshold is when these two values are exactly equal. As will become clear, this can occur only when voter tolerance is high and the intervals of party support meet in the middle of the distribution. We denote this threshold by $\hat{\lambda}(\hat{r})$, where

$$\int_0^{\hat{r} + \hat{\lambda}(\hat{r})} v f(v; \alpha) dv = \hat{r}.$$

We then have the following result.

Proposition 3. *Fixing party positions $\hat{r} = -\hat{d} > 0$, polarization decreases monotonically if $\hat{r} < 2\alpha \ln 2$ and $\lambda > \hat{\lambda}(\hat{r}) > \hat{r}$, otherwise it increases monotonically.*

To understand why the impact on polarization is ambiguous, begin with the case in which voter tolerance is low and abstainers exist in the middle of the electorate. This implies the intervals of support are symmetric around each party's positions. Because the logistic distribution is single-peaked, there are more voters on the inside of a party's position than on the outside. Mechanically, therefore, updating causes more outward than inward movement and the aggregate effect is to increase polarization.

For higher levels of voter tolerance this effect can reverse. For higher levels of tolerance the intervals of support for the parties intersect and compress in the center, potentially leaving more voters on the outside flank of each party. Once this compression is sufficient to push the average ideal point to the outside of the party position, the effect on polarization reverses and the citizenry becomes more moderate, and less polarized, over time.

The average ideal point of voters is important because this average will inevitably converge over time to the party position. Intuitively, as voters collapse in on the party position, so too must their average. Mechanically, if the average begins outside the party position, the aggregate effect will be for moderation. [Proposition 3](#) establishes that this average is all that matters.

The critical thresholds of $2\alpha \ln 2$ and $\hat{\lambda}(\hat{r})$ demarcate the point at which the crossover occurs. The value of $2\alpha \ln 2$ is the initial average ideal point for all citizens to the right of 0, whether voters or abstainers. As the logistic distribution is single-peaked, the average of voters must be inside this threshold when abstention is only on the flanks. Thus, if party R is more extreme than $2\alpha \ln 2$, the initial average ideal point of voters must be to its inside and the aggregate dynamic is for polarization.

For more moderate party positions, the average ideal point of voters can initially be more extreme than the party. This is where the second threshold, $\hat{\lambda}(\hat{r})$, comes in. As noted, this crossover can only occur when $\lambda > \hat{r}$ and the compression of a party's

support at the center is sufficient. An interesting case emerges when party R is located between $2\alpha \ln 2$ and the median citizen on the right side of the distribution at $\alpha \ln 3$ (that the median is more moderate than the average follows from single-peakedness of the logistic distribution). In this case, more voters are on the inside of the party position than outside, and thus more voters are shifting outward in their ideal point than are shifting inward. Nevertheless, in aggregate, the electorate is moderating because the voters on the outside are more extreme than the moderate voters are moderate, and the moderation of extreme voters outweighs the polarization of moderate voters.¹⁴

Voter updating with fixed party positions leads to rich dynamics but, ultimately, it can only explain so much. On their own voters may polarize, but they may also moderate. Even if they polarize, the effect is bounded by the locations of the parties initial policies. Moreover, the preference profiles that do emerge are inconsistent with other known properties of voting, such as negative partisanship.¹⁵ On top of this, there is, of course, no polarization of the parties. In the following section we reintroduce strategic parties and show how their reactions to voter updating creates an interdependence and co-evolution of elite and mass positions that does resonate with the data.

5 Adding Strategic Parties Back In

Strategic parties respond to the changing distribution of ideal points, which, in turn, changes the evolution of voter preferences. In this section we characterize the dynamic process that results, beginning with the second election. Throughout we presume that the policy positions in the first election are those given by the unique symmetric equilibrium described in [Proposition 1](#).

5.1 The Second Election

The gaps in the distribution of ideal points fundamentally change the incentives of the parties. The equilibrium positions in the first election balance the incentive to converge

¹⁴The divide that forms between voters and abstainers, and the growing homogenization of voters, resonates with the axiomatic measures of polarization of [Esteban and Ray \(1994, 2012\)](#). A key difference is that the compression of voters here is around the location of each party rather than the mean of the group distribution as it is in Esteban and Ray's notion of a "squeeze." Differences aside, the growing divide between voters and abstainers in U.S. politics increasingly resonates with the in-group and out-group measures in Esteban and Ray's work, suggesting that their theory of polarization that was developed in the context of conflict and ethnic and tribal allegiance, could also find profitable application in the domain of U.S. politics. ([Clark \(2009\)](#) offers the one application of their ideas to the U.S. in the context of Supreme Court justices.)

¹⁵All voters experience increasing preference for their favored party, whereas negative partisanship finds that this is relatively stable.

to gain more votes against the incentive to diverge to a better policy position. Starting from the same positions, that trade-off now collapses. The gap(s) in the center of the distribution imply that a party can diverge slightly without losing any votes. No votes are lost because there are no voters there to lose. Those that had been on the inside margin and who did turn out to vote updated toward the party, leaving behind an empty space. This changes the calculus of the parties and they respond by moving their positions toward the extremes.

This logic provides the foundation for polarization. It is, however, only half of it. As one party shifts outward, so does the other party, and, as a result, the mid-point between them remains unchanged. This implies the incentive to shift outward is recreated anew, leading to more polarization and potentially a substantial unwinding of party positions. This unwinding replicates but turns on its head the classic logic of convergence due to [Hotelling \(1929\)](#). In the classic intuition, parties inch toward the center to win the median voter and, as the opposition does the same, this creates an iterative process, that leads to full convergence. In our model, in contrast, a party inches outward without losing votes, the opposition party responds, and the iterative process leads instead to polarization.

Unlike in Hotelling, however, the iterative process need not lead to complete unraveling instantaneously. The parties in our model may no longer be constrained by each other, but that does not allow them to escape from competition altogether. Instead of competing against each other, the parties compete against voter apathy. If they polarize too much, the parties will lose voters to abstention.

The exact nature of that divergence depends on the type of the first election equilibrium and, thus, on the level of voter tolerance. For low levels of voter tolerance, i.e., $\lambda \leq \lambda^*$, the parties do not compete directly in the first election and their ability to polarize is limited only by the extent of their own voters' updating. The largest shift in preferences is by the voters who were on the margin between voting and abstaining in the first election. [Proposition 4](#) shows that this amount, $\lambda\tau$, is exactly the amount that the parties polarize at the second election.

Proposition 4. *At the second election, the unique equilibrium for $\lambda \leq \lambda^*$ is:*

$$\begin{aligned} r_2^* &= \min \{r_1^* + \lambda\tau, R\} \\ d_2^* &= \max \{d_1^* - \lambda\tau, D\}. \end{aligned}$$

The equilibrium represents a sort of “no voter left behind” strategy. The parties only polarize as much as they can without losing any voters to abstention. Any larger polarization and the marginal voter on the inside would roll off and abstain, even allowing for that voter's own outward shift.

While the parties are leaving no voters behind on the inside, they are also gaining voters on the flank. Abstainers on the flank who were just outside the margin of voting

in the first election are now $\lambda\tau$ closer to the new party position, and as a result, an interval of abstainers of that length switch to voting and turnout goes up.

The updating of voter preferences allows the parties to, in a sense, secure their core supporters and this, in turn, gives the parties freedom to move. Rather than move to the center to appeal to abstaining moderates, however, the parties use the opportunity to polarize outward, drawing more extreme citizens into the voting pool. This implies that as the voting pool grows, it is the newer voters that are the most extreme.

The situation when voter tolerance is high ($\lambda > \lambda^*$) leads to even more polarization in the second election, although it can also lead to the equilibrium failing. For high voter tolerance, the intervals of support in the first election intersect, which implies that the length of party support on its inside is less than the full length of λ . The shortened length means less updating by voters, with the marginal supporter for party R at 0 only moving her ideal point outward by $r_1^*\tau < \lambda\tau$. This might suggest that the freedom of the parties to shift outward is also compressed, but, in fact, the opposite is true and the parties polarize to a greater extent.

The increased freedom to polarize comes from the fact that the compressed intervals of support represent slack in the parties' ability to win voters. In the first election the parties win only an interval of support of length r_1^* to their inside whereas voter tolerance is λ , meaning there is $\lambda - r_1^*$ in slack that can be exploited. To put it another way, the inside boundary of party support does not so quickly hit its limit as the parties shift outward. Combining this with the fact from above that competition is against voter apathy rather than the other party directly, slackness allows the parties to polarize faster. To implement the "no voter left behind" strategy, therefore, party R can at most leave the voter located at $r_1^*\tau$ indifferent, which translates to a location for the party at $r_1^*\tau + \lambda$. This can represent a substantial jump from the first election position when λ is large, and potentially all the way to the parties' ideal points. [Proposition 5](#) confirms that this is, indeed, the equilibrium for $\lambda \geq \lambda^*$, with polarization bounded by the parties' own ideal points.

Proposition 5. *At the second election, there exist $\bar{\lambda} > \lambda^*$ and $\bar{\tau} \in (0,1)$ such that for all $\lambda^* < \lambda \leq \bar{\lambda}$ and $\bar{\tau} \leq \tau < 1$, the unique equilibrium is:*

$$\begin{aligned} r_2^* &= \min \{r_1^*\tau + \lambda, R\} \\ d_2^* &= \max \{-|d_1^*|\tau - \lambda, D\}. \end{aligned}$$

For $\lambda > \bar{\lambda}$ and $\tau < \bar{\tau}$, there exists no symmetric equilibrium.

The equilibrium implies that even small changes in the distribution of voter ideal points can lead to substantial and immediate polarization of the parties. This is because voter updating of any size causes a gap to open up in the distribution, and it is this

gap that induces tit-for-tat divergence such that the parties unwind their positions to the point where they are no longer competing against each other but against voter apathy and abstention.

The logic of the result does have a limit as the equilibrium fails for sufficiently high λ and small τ . This failure derives from failure of the second order condition: eventually, as polarization increases, a point is reached at which the parties find it profitable to deviate and jump to the center. Failure occurs because, in effect, the logic of divergence is too powerful and the parties otherwise get too far apart too quickly. To see this, the distribution of ideal points in the second election when τ is small is very close to that in the first election. The leave-no-voter-behind strategy, however, can generate substantial polarization. This leaves a large block of voters in the middle who can be exploited, and, eventually, one of the parties prefers to do so. This does not mean, though, that an equilibrium exists with centrist positions as then the same unwinding logic would again apply.¹⁶

The most striking feature of the equilibrium in [Proposition 5](#) is that the type of electoral competition changes. Electoral competition in the first election for high λ is of the classic win-the-median-voter form, yet by the second election, this style of competition has given way to a mobilize-the-base strategy. Therefore, by the second election, electoral competition is such that the parties do not compete head-to-head for the median voter, but instead compete only indirectly, focusing on the margin of turnout rather than persuasion. This pattern continues through later elections, suggesting that rather than being the unusual case, this style of competition is the norm.

The leave-no-voter-behind strategy is intuitive yet seeing exactly why it is optimal requires some digging. It is clear that if the parties polarize, they should polarize no less than they do with this strategy. Polarizing less would strictly decrease their vote share and implement a less appealing policy. What is less clear is why the parties do not polarize further, in fact leaving some voters behind, or why they do not instead exploit the opportunities created by the gaps in the distribution to converge. The answer to both questions comes from the logic of the first period equilibrium.

The first election equilibrium tells us that it is not profitable for a party to deviate outward as the loss of centrist voters outweighs the gain in voters on the flank and the more appealing policy position. That the parties do not wish to polarize more than with the no-voter-left behind strategy follows from this by a simple dominance argument. At its new position, the inside flank consists of exactly the same voters as in the first election (as they updated by $\lambda\tau$ in [Proposition 4](#) and by $r_1^*\tau$ in [Proposition 5](#)), although now these voters are packed more densely and the rate of loss from further divergence is higher. At the same time, the marginal voters that would be gained on the flank are fewer in

¹⁶For an equilibrium to exist in this situation it would need to be asymmetric or in mixed strategies. We leave the nature of this equilibrium (or whether one exists) as an open question.

number (lower density further out), and the policy cost of losing is now higher as the opposition party has shifted further away. Consequently, if deviating outward from r_1^* in the first period is not profitable, deviating outward from r_2^* in the second election is also not profitable.

Typically, this strict dominance argument would imply that the party must then find it optimal to instead shift inward. This would be true if expected utility were continuous. However, because of the gaps in the distribution of ideal points, expected utility is discontinuous in location (in contrast to the first election) and inward deviations are not profitable. Putting the two pieces together ensures the strategies in [Proposition 4](#) and [Proposition 5](#) constitute local optima.

Surprisingly, the same logic can also be used to support a second local optimum. This is the case when voter tolerance is low and centrist citizens abstain in the first election ([Proposition 4](#)). The logic of the no-voter-left-behind strategy works also in reverse and the parties converge to the center rather than polarize. In the same way as for divergence, converging does not lose voters on the flank (who have updated inward) whereas it gains voters in the center. We can use the fact that the parties are indifferent about converging from the equilibrium location in the first election to show that they strictly prefer to converge from the same position at the second election, even though converging delivers a less appealing policy should the party win.

This possibility complicates the analysis as expected utility is no longer quasiconcave. This makes it difficult to establish that the strategy in [Proposition 4](#) is not only a local optimum but also a global optimum—and, therefore, an equilibrium—but also that a second, more convergent equilibrium does not exist. In the proof in the appendix we construct a dominance argument that shows that convergence is dominated by polarization, thereby accounting for both of these concerns and establishing our equilibrium result. The main intuition boils down to the relative weight of the voters who update outward versus those who update inward. Because more voters are on the inside rather than outside of the party position, the relative gain of no-voter-left-behind is greater for polarization than for moderation.¹⁷ On top of this, the opposition party's voters, who have also updated their ideal points, are further away and harder to capture, and the relative cost of losing is higher when the opponent polarizes rather than moderates. By putting these pieces together, combined with the fact that jumping to the center is not profitable in the first election, we establish that no-voter-left behind *with* polarization is the unique equilibrium.

¹⁷In the first election, convergence from the equilibrium may lose x voters on the flank and gain $y > x$ voters at the center. From this same location, the gains in either direction in the second election are the same whereas diverging now avoids losing y voters whereas converging avoids losing only x voters. This implies divergence is relatively more attractive.

5.2 The Third and Subsequent Elections

Voters update their ideal points again after the second election, the parties respond and the process iterates. As this process continues the parties progressively polarize, continuing until they reach their ideal points, at which they stabilize. This can take a few elections or it can take many, depending on the party ideal points themselves as well as the speed at which voters update their ideal points and follow the parties' positions.

A difference at the third election and thereafter is that the nature of the equilibrium no longer depends on the size of λ . In the second election the parties polarize such that the marginal voter is λ from the party position. The “no voter left behind” logic then implies that the parties can polarize exactly $\lambda\tau$ further in each election. The recursive process of polarization that this sets in motion is described in [Proposition 6](#).

Proposition 6. *For election $t \geq 3$, the unique equilibrium is symmetric and as follows:*

$$\begin{aligned} r_t^* &= \min \{r_{t-1}^* + \lambda\tau, R\} \\ d_t^* &= \max \{d_{t-1}^* - \lambda\tau, D\}. \end{aligned}$$

Figure 5 shows the polarization process for different values of $\lambda\tau$. The rate of polarization is constant in each case with the exception of panel (c) for large voter tolerance. In this case a kink appears at the second election as the parties exploit their latent appeal to voters before settling down at rate $\lambda\tau$. The faster start in case (c) does not necessarily imply faster polarization overall. Should τ be small such that $\lambda\tau$ is also small then, as depicted in panel (c), polarization is thereafter slow and drawn out.¹⁸

The polarization of the parties leads to the polarization of voters, although the pattern of voter polarization is more varied and the timing different. We emphasize four features of voter behavior that stand out.

The first notable feature of voter polarization is that in the long run it is unambiguous. If the parties polarize then so too do voters. This differs from the case of fixed party positions in which voter polarization is ambiguous ([Proposition 2](#)), confirming that it is the interaction of party positions and voter updating that drives voter polarization.

The second notable feature of voter polarization is that it need not be monotonic, both at the aggregate level and for individual voters. This also differs from the situation with fixed party positions, and is also different to the polarization trajectory of the parties. [Figure 6](#) depicts the possible paths of voter ideal points for supporters of party R .

The trajectory of ideal points is monotonic for two sets of voters, the most moderate and the most extreme, although the direction of movement is the opposite for the two groups. Voters who are initially more moderate than the parties polarize monotonically,

¹⁸The starting party positions, d_1^* and r_1^* , can be calibrated by varying the parameters α , β , and λ .

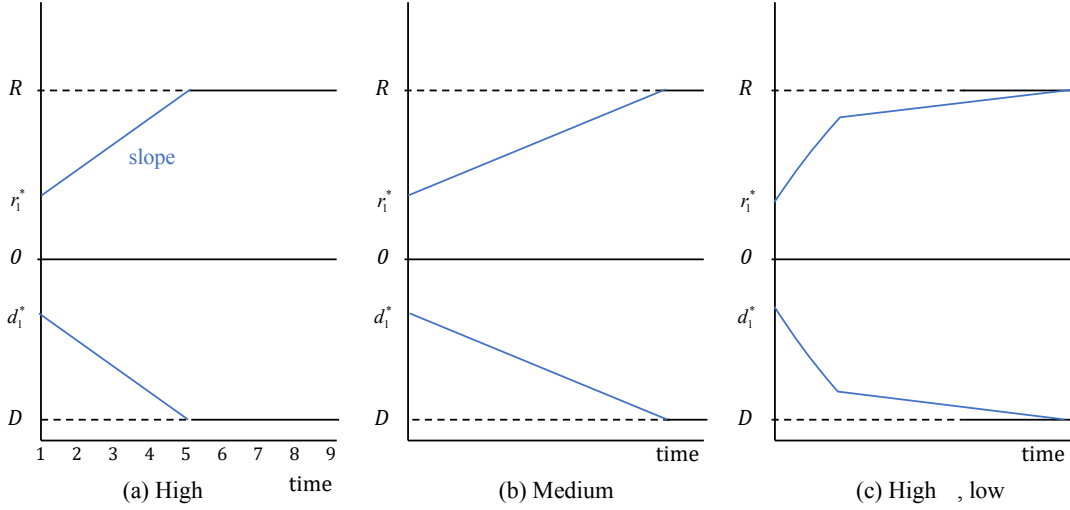


Figure 5: Party Polarization Over Time

chasing, in effect, their favored party as it moves outward. In contrast, the voters who are most extreme—with ideal points beyond the party ideal points—begin by abstaining but are eventually drawn in off the sidelines to vote, and when they are, they traverse a monotonic path inward. Both sets of voters ultimately converge on their preferred party's ideal points

The path of polarization is non-monotonic for voters between these two extremes. A citizen with ideal point between a party's ideal point and first election position will begin to moderate her position once drawn in to vote. Eventually, however, the party will cross over her ideal point, causing that voter to reverse course and begin to polarize outward. Ultimately, this voter will polarize past their own initial ideal point, continuing outward until the party stops at its ideal point and the voter begins to catch up.¹⁹

A third notable feature of voter polarization is that it lags that of the parties. Interestingly, the parties not only polarize ahead of most voters, they also polarize faster. At each election, each party polarizes by an amount $\lambda\tau$ (and by further in the second election when λ is large), whereas the voters polarize strictly less than this amount. Only precisely at the boundary of a party's support—the voter exactly λ from the party's position—is updating by the full amount of $\lambda\tau$. All voters with ideal points closer to the party update by less. Consequently, the parties actually move away from their more moderate supporters during the polarization process. Only when the parties stabilize at their ideal points do the more moderate voters begin to catch up.

¹⁹Citizens with ideal points only just beyond the party ideal point will exhibit a mixture of these properties. When drawn in to vote, they may cross over R . Thus, when the party position crosses their ideal point they will reverse course and polarize, but they will converge only on R and not again reach their own original ideal point.

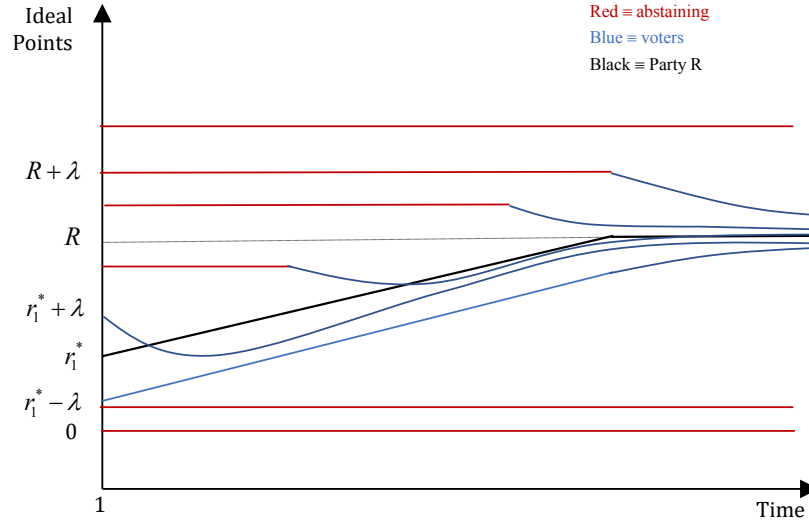


Figure 6: Path of Citizen Ideal Points

The slower polarization of voters manifests at the aggregate level in the distribution of ideal points that develops. As the parties polarize, they pass many of their own voters, which leads to a large majority of voters being located on the inside of each party's position. Moreover, the faster polarization of the parties leads to much of this mass accumulating at the inside fringe of each party's support. Thus, the great bulk of a party's support ends up being more moderate than the party itself, reinforcing the impression that voters lag the parties as they polarize.

Figure 7 depicts the distribution of citizen ideal points on the right side of 0 after the election at which party R first locates at its ideal point. The coarseness of the figure obscures much of the richness of the distribution. At a fine micro level, the distribution has both lumps and discontinuities. On the right side of party R , each new interval of citizens drawn in to vote create their own voting block disconnected from the other voters, with the gaps contracting over time. On the left side of R , the distribution has no gaps but it has lumps as these cohorts are drawn in to vote and become embedded into the overall distribution. Of course, this distribution is itself only transitory as, even though the parties no longer change positions, the voters continue to converge on the parties and the distribution of ideal points increasingly collapses around these points.

The fourth and final feature of voter behavior we emphasize is turnout. As the parties polarize, the "no voter left behind" strategy means no voters are ever lost to abstention, whereas new voters are gathered in on the flanks (as is evident in Figure 6). Thus, aggregate turnout strictly increases from election to election until the parties reach their ideal points, after which turnout stabilizes. Turnout will remain incomplete as abstainers out on the far flanks will never be drawn in to vote and, for low levels of voter tolerance, centrist abstainers will only grow ever more alienated. Figure 8 depicts the limit levels

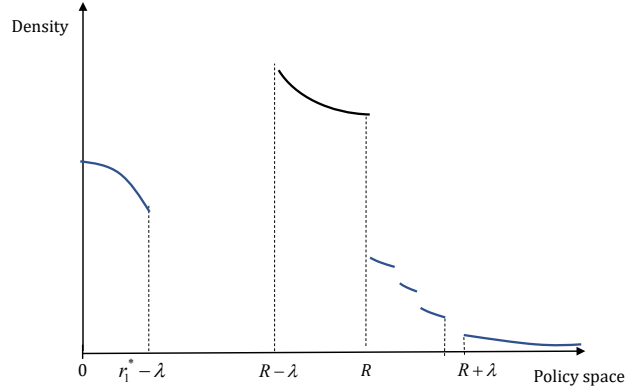


Figure 7: Distribution of Citizen Ideal Points

of turnout overlaid on the turnout rate at the first election for $\beta = 0$ depicted earlier in Figure 3. An interesting effect is that the turnout gap between low and high values of λ contracts over time, both in an absolute sense and more dramatically in a relative sense. This is because, regardless of the level of voter tolerance, the parties end up at the same policy position and sweep up all of the voters in their path as they traverse their path outward.

5.3 Discussion

The features of polarization described in the previous section, with the exception of turnout, resonate with the data. Polarization of elites has been significant, and it occurred earlier, faster, and to a greater extent than polarization of voters. The model also rationalizes the ostensibly distinct phenomenon of negative partisanship. All voters are moving away from the opposing party as the parties polarize. More strikingly, at the early stages of the polarization process most voters are not getting closer to their favored party, and many are actually falling further away. It is the distinct combination of candidate evaluations that define negative partisanship.

The dynamic path of turnout does not fit the data as well. Although it is common to lament a decline in turnout in U.S. national elections, the evidence suggests that turnout has been relatively constant throughout the era of polarization (McDonald and Popkin, 2001). Either way, it does not match the prediction of increasing turnout in the model. In the following section we offer an extension to the model to allow for generational turnover, and we show how this realistic enrichment reconciles the model's prediction with empirical observation. Indeed, it explains non-increasing turnout in a way that also matches the differential patterns of turnout across generations.

In addition to explaining the past, we can also put the model to work of prediction. What is the future of polarization? Will elites and masses continue to polarize? Will

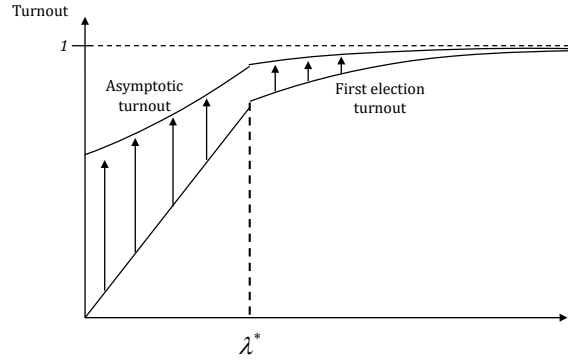


Figure 8: Dynamic Turnout

elites always be more polarized than the masses? Our model suggests answers to these questions. According to our model, the parties continue to polarize until they reach their own ideal points. Only then will polarization of elites stop. The polarization of the masses will continue beyond this time, albeit at a slower rate as voters converge upon the party location. The end point of the model is for voters of each party to form an homogeneous block at the location as their favored party and removed from non-voters. That in practice voters are currently less polarized than the parties suggests this process has not yet run its course, surely discomfoting news for those who already lament the polarized state of politics today.²⁰

The dynamic just described also predicts that negative partisanship will ultimately weaken. In the final convergent phase, most voters will continue to increase in their dislike of the opposition, although some will begin to rate the opposition more favorably (the most extreme voters), whereas all voters will begin to rate their favored party more highly.

The model also speaks to an ongoing debate in the political science literature on the extent to which the masses have polarized, if at all. At the heart of this debate are two conflicting pieces of evidence. On one hand, support for moderate positions remains large, even after decades of elite polarization (Fiorina et al., 2010). On the other hand, voters, and particularly those most engaged in politics, have polarized considerably (Abramowitz, 2010). The patterns of behavior in our model offer one way to adjudicate this dispute. Although the mechanism in our model is simple, the dynamic shows how rich and varied voter behavior of this sort can emerge from the simple process of elite polarization. In our model, low voter tolerance implies the existence of centrist abstainers, and over time, these citizens remain exactly where they are. This rationalizes the evidence that moderate policies remain supported. At the same time, those who

²⁰We do not consider third party entry in our model, although this evolution points to a simple explanation for why entry does not occur—there are no voters in the middle!—despite the polarization of the two major parties.

choose to vote polarize and, consistent with [Abramowitz \(2010\)](#), the polarizers are those most engaged in politics, which creates a bimodal distribution of ideal points.²¹

More broadly, our model connects to the broader debates in political economy about the origin and nature of political preferences. On one side is the classic spatial theory of voting familiar from political economy. This theory places ideology and policy at its center, with a measure of distance that determines vote choice and political outcomes. In this theory, what [Hall and Thompson \(2018\)](#) label the “institutional literature,” voters can swing from one party to the other if it moves closer. Opposing this view is what [Hall and Thompson \(2018\)](#) label the “behavioral literature.” In this literature, dating back to [Campbell et al. \(1960\)](#) and [Converse \(1964\)](#), ideology plays little role and swing voters don’t exist. Instead, voters are rigid partisans who are rallied to their team at each election. According to this view, voting is a purely partisan endeavor and ideological preferences are nothing more than ex post rationalizations of behavior. The model we introduce demonstrates how these contrasting perspectives, and the evidence in support of each, can be unified. Our theory is very much in Hall and Thompson’s “institutional literature.” Yet by endogenizing the ideal points of citizens, we demonstrate how patterns are generated that resonate with the findings of the behaviorists. We show how voting is spatial within elections, such as found by [Jessee \(2009, 2010\)](#), yet at the same time for preferences to be responsive to party cues.

This duality can also inform the puzzle of why parties do not choose more moderate positions to appeal to the median voter. The behaviorists interpret this absence as evidence that voting is not spatial and ideological. Our theory suggests the potential error in this inference. In our model, voting is purely ideological, but due to the updating process and the polarization of the parties, a party relocating to the middle is not profitable as the middle has been hollowed out, with the unattractiveness of this strategy increasing the longer the polarization process goes on. The “disappearing center,” as [Abramowitz \(2010\)](#) calls it, implies that even in a purely ideological world, the distribution of citizen ideal points can become bimodal, and political competition can resemble a contest of voter mobilization rather than one of persuading the median voter, as evidence suggests has become the dominant strategy amongst parties in U.S. elections.

These broader debates have previously found their way into the literature on polarization. The behaviorist literature, in dismissing ideology and voter agency, naturally converged on the conclusion that polarization is elite-led ([Achen and Bartels, 2016](#); [Lenz, 2012](#)). As neatly as this behaviorist view rationalizes polarization, it fails to explain why polarization has been slow and iterative rather than dramatic, and indeed, why there needed to be a process at all—if elites are unconstrained ideologically, why were they ever *not* polarized? These gaps demonstrate the importance of explaining both the

²¹Updating in our model distinguishes only between voters and abstainers. A generalization, which may more closely match the data, is to calibrate updating to the degree of a citizen’s engagement in politics.

state and the *process* of polarization. We match the data that the behaviorist literature emphasizes—that it is the elites that polarize first—but, additionally, we are also able to explain both the progressivity of polarization and the sequencing of polarization by the elites and the masses. Critically, the rationale for polarization that we offer is entirely at odds with the behaviorist view of politics. We show that what might appear to be an elite-driven polarization process, is actually a process driven by voters and the manner in which their ideological preferences are constructed.

5.4 Asymmetric Polarization

The power of our model is in explaining rich dynamics with a parsimonious specification. Not surprisingly, there is much that remains beyond the model’s reach. One such feature of particular prominence is that polarization has in practice been asymmetric, with the Republican party polarizing more than Democrats. As with polarization itself, this asymmetry has been both in the speed of polarization and, at least till now, the extent of Republican polarization.

That polarization in our model is symmetric follows directly from the symmetry of the model’s construction. Capturing asymmetric polarization requires relaxing symmetry in some way. In this section we offer three natural extensions to the model that can accommodate both the process and the final state of asymmetric polarization, as well as other features of political competition that resonate with the data. In choosing these extensions we have focused on the feedback loop between voters and parties, as we have done in the model so far, showing how micro-asymmetries in voter behavior can reverberate through the system to produce macro asymmetries in party positioning.

An obvious possibility why Republicans have polarized more than Democrats is, simply, that they possess more extreme preferences. [Corollary 1](#) describes the equilibrium in this case. It is a straightforward extension of the equilibrium in the baseline model, indeed, it is exactly the same up to the point at which party *D* reaches its ideal point. Once this is reached, *D* stabilizes whereas party *R* continues to polarize further until it reaches its own ideal point. Asymmetric polarization manifests, therefore, as a longer polarization process for Republicans rather than a faster one. The equilibrium follows directly from the baseline model due to the linearity of utility and complete discounting. To save on excessive notation, we state the results somewhat informally.

Corollary 1. *Suppose $R > -D$. The equilibrium is the same as described in Propositions 1-6 up to election t' when $d_{t'}^* = D$. For elections $t > t'$, $d_t^* = D$ and $r_t^* = r_{t-1}^* + \lambda\tau$ until it reaches position R , after which it too stabilizes. For each $t \geq t'$, the probability *R* wins the election is greater than 50% and it is strictly increasing until *R* reaches its ideal point.*

The striking feature of the equilibrium is that not only is party *R* not punished for its greater polarization but that it actually benefits from it. The fundamental intuition from

spatial models of politics is that the more centrist party wins election more frequently. The reason for this contrast—for party R 's greater electoral success—is that the process of polarization matters and not just the state. Over the process of polarization, parties carry their voters with them to the extremes, sweeping up all of those in their path. Counterintuitively, therefore, the more a party polarizes, the more abstainers it pulls into its orbit and the higher its vote share. This fact offers a rationalization for why the Republican party has still managed to win elections not despite its greater polarization but rather because of it.

Asymmetric polarization may also derive from the voters rather than the parties, for example when the speed of updating is party specific. Suppose that right-wing voters update their ideal points more fervently after voting (so $\tau_R > \tau_D$). They may do so because Republican voters have developed a stronger partisan identity or have more partisan news consumption (i.e., Fox news). Regardless of source, stronger updating by Republican voters allows the Republican party more freedom to polarize and, in this case, to not only polarize more but also to polarize faster than the Democrats do.

Corollary 2. *Suppose $0 < \tau_D < \tau_R < 1$. The equilibrium follows the pattern of Propositions 1-6 except $r_t^* > -d_t^*$ at each election until election t' when $d_{t'}^* = D$. The parties win with equal probability at the first election and at election t' and on. Between elections 1 and t' , R wins the election with greater than 50% probability.*

Party R again benefits electorally from its greater polarization. Its faster polarization allows it to pull in more voters more quickly, giving it a greater probability of winning as well as a more attractive policy to implement. This advantage persists as long as R is more polarized than party D , and disappears once D catches up and both parties stabilize at their (symmetric) ideal points. This contrasts with the asymmetry evident in [Corollary 1](#) that is slower to arrive but permanent when it does. It is possible, of course, that both of these asymmetries are present in practice, which would generate both an early and persistent competitive advantage for party R .

A third possibility is that asymmetric polarization is the result of random chance. [Beasley and Joslyn \(2001\)](#) present evidence that the feedback loop from action to preferences depends on whether a candidate wins election or not, with supporters of the winner updating more than supporters of the loser. Even in the symmetric set-up, therefore, the random draw of the election winner can endogenously create an asymmetry.

The trajectory in this setting is path dependent. An early lucky win causes a party's voters to update more, giving it greater freedom to polarize and draw in more abstainers from the flank, which, in turn, gives it a greater probability of winning the next election. This reinforcement can create momentum effects that allows an early lucky break to be built into a sustainable advantage. Lucky breaks can run both ways, however, so unlike [Corollary 2](#), an early advantage here is likely to persist but it can disappear.

Momentum effects and path dependence are ubiquitous features of politics (Pierson, 2000, 2004; Page, 2006). In the context of asymmetric polarization, this suggests that the greater Republican polarization may be due to an early lucky break rather than innate differences, perhaps even due to the good fortune of charismatic Ronald Reagan arriving on the scene as the era of polarization was building up steam. Unfortunately, this extension of the model is more analytically complicated than the previous two, and we do not move beyond sketching the intuition here.

6 Concluding Discussion

To illuminate the feedback loop between parties and voters we have sought to keep the model as simple and transparent as possible. The results are robust to many of these choices. Two are worthy of closer inspection. One relates to the foresight of leaders. A strength of our result is that it emerges even with parties that aren't forward-looking, and, therefore, doesn't rely on any sort of scheming or inter-temporal trade-offs. Iterative polarization comes principally from the feedback loop between voters and parties. If the parties were forward-looking, the feedback loop would continue and the polarization would only be more dramatic. A straightforward intuition is that the intensity of competition in the first election would increase as the parties foresee that a higher vote share initially will carry over to future elections. This is akin to the impact of switching costs in markets. Klemperer (1987) shows how price competition is more intense early and weakens later once consumers are attached to one supplier.²² The same dynamic is likely to play out here: more convergence in the first period as parties compete intensely for votes but with the same end-point (full polarization), creating an even broader sweep of polarization by the parties over time.²³

A second open question regards the preferences of the parties. In our model the party ideal points are fixed. The implication is that parties have always been extreme and that polarization has allowed them to express those preferences in policy. In practice, the preferences of political parties are a complicated amalgamation of many forces,

²²An even more explicit strategy was laid out by the Chinese philosopher, Laozi, and how a leader can exploit the attachment of the people to serve his own ends: "The ruler is thus able to accomplish everything, but it will seem to the people as though everything is simply occurring naturally, without any directing will: 'When his achievements are completed and tasks finished, the commoners say that 'We are like this naturally (*zi ran*).' " (Puett and Gross-Loh, 2016)

²³Allowing for forward looking behavior suggests some interesting possibilities. Anticipating the Republican's long-term advantage, Democrats may intensify first-period competition and initially converge further, seeking to win more voters that it can then drag to the extreme with it. A more novel possibility emerges at the other end of the polarization process. It may be that the parties polarize *beyond* their own ideal point to gather more voters, then reversing course and settling back to its ideal point. (This possibility may even occur with myopic parties and convex utility as the competitive margin of winning more votes actually pushes parties outward rather than inward.)

and it is plausible, indeed likely, that the political parties have themselves developed more extreme preferences over time. A natural conjecture is that party preferences are an aggregation of the preferences of their members or, more narrowly, their elected representatives. Our results would carry over directly to such an environment as long as the preferences of those within the party evolve in a way that is more polarized than among the electorate as a whole.²⁴ In this richer model, the feedback loop would move from voter preferences to the preferences and membership of the parties, and then into the party platform, before closing the loop to voter preferences. Documenting and understanding this mechanism rigorously is a promising direction for further work.

Our model opens a new perspective on political representation. If the preferences of the citizenry are evolving, what does it mean for them to be represented in politics? Should representation be measured by where citizens begin, where they end, or by a series of snapshots at each point along the dynamic process? Interestingly, the updating of preferences by voters embeds a positive force for representation into the system, as, at least measured naively, representation inexorably increases as voters and parties converge. In contrast to standard notions of representation, however, the convergence is not because parties move to where voters are, but because voters move to the parties. The problem of political representation raised here relates to the problems in measuring welfare in behavioral economics (Bernheim and Rangel, 2009). Whether those tools can be extended to political representation, or whether a new framework is required, is an important open question.

The focus of our paper is on the era of polarization that began in the second half of the 20th century and has run through to the present day and, presumably, into the future. The evolution of political preferences did not begin with this era, of course, and prior to the era of polarization was an extended period of convergence. In fact, the major parties in the U.S. had reached such a point of convergence in the 1950's that the American Political Science Association famously issued a report lamenting that the parties were too close together and that they offered voters an insufficiently differentiated choice. This juxtaposition leads to the question of why then. Why did polarization begin when it did? And what caused the moderation of parties in the earlier era? An explanation of the broader sweep of American political history is, unfortunately, beyond the scope of the present paper, although obtaining such an explanation is of clear importance. Toward this end, and to show how our theory of polarization can be used as a piece of this broader understanding, we sketch here how our model can be extended to capture an even longer and richer period of preference dynamic.

One restriction of our model is that the population of citizens is fixed. This is particularly important as we seek to explain a dynamic phenomenon that stretches over more than half a century, and during that time there has been substantial turnover in

²⁴The data supports this relative ordering of voters and party members all throughout the period of polarization (Abramowitz, 2010).

the citizens of voting age. Consider then our model with the addition of births, deaths, and generational turnover. To be more specific, suppose that after each election a new generation is born of mass ρ and that each generation lives for some finite number of periods, T .²⁵ To fix ideas, suppose that each generation arrives according to the same original distribution F .

The population at the first election will be the same as in the baseline model and the equilibrium will go through unchanged. For the second election the equilibrium logic continues to go through as long as the mass of the new generations born isn't too large, and the same holds for the third election, and so on.²⁶ The parties begin the same polarization process and, as in the baseline model, the same patterns of elite and mass polarization begin to emerge.

Fast forward then to the point where the parties reach their own ideal points and stabilize. In the baseline model no new voters are won over and the set of existing voters remains stable, converging in preference around the parties. With births and deaths, however, this set of voters is not stable, but is dying off approximately at rate ρ . At the same time, the new generations are arriving with distribution F , and if the parties are polarized, not many of these new voters are falling into each party's interval of support. This leads to a different trajectory for turnout than in the baseline model. Rather than increasing throughout the polarizing phase and remaining stable thereafter, turnout is tempered by the newly born who aren't captured by one of the parties, and it declines once the parties stabilize at their ideal points. This pattern matches more accurately the evidence from U.S. elections (McDonald and Popkin, 2001), reconciling a discordant prediction from the baseline model.

The changing pattern of turnout matters also for polarization and the location of the parties. As a large mass of unattached citizens accumulates in the center of the distribution, the opportunity emerges for a party to move their position and appeal to them. It is intuitive to see that this generational turnover will lead inevitably to the end of polarization. Eventually one of the parties will find it optimal to move to the center to win over abstainers. An open question is whether the movement inward will come suddenly or progressively. Will a party make a sudden jump to the center, alienating their core supporters, whilst awakening a new generations of voters? Or will the parties inch inward progressively attracting new voters and dragging their old voters with them toward the center? Whichever is the answer, the inevitable lure of the center will surely comfort those concerned with today's state of polarization, yet each possibility implies very different timing for the end of polarization and suggest different types of politics.

The inclusion of generational turnover provides a natural and simple explanation

²⁵To allow a seeding process, we can think of the original generation dying off at rate ρ from the second election onward.

²⁶As this is only a sketch, we do not dwell on the precise range of ρ that satisfy this argument, although it is clear it is not empty.

for cycles of moderation and convergence. Appealingly, it also creates patterns at the micro level consistent with observation. The assumption that each generation arrives according to the same distribution, F , presumes that citizens come of age unmolded by politics. Although somewhat extreme, this assumption is nevertheless consistent with the core assumption of our model that the act of voting itself shapes political preferences. Empirically, it is supported by evidence that a citizen's political preferences are shaped significantly by the first presidential election in which they are eligible to vote. Indeed, [Ghitza and Gelman \(2014\)](#) show how a citizen's lifetime of presidential election shapes their preferences in a sort-of running tally way, with by far the most weight on the first election.²⁷

Regardless of how polarization breaks down and party moderation obtains, it is clear that the parties, having captured the mass of young centrist voters, will once again discover the incentive to polarize, and the polarization process will begin anew. Of course, throughout the period of polarization, new generations will continue to be born, and with these new generations arriving uncaptured at the center, the process of polarization will, eventually, break down again. The cycle that this creates demonstrates how the moderation of the first half of the 20th century can fit naturally with the polarization of the second half.²⁸ It also suggests why the alarm of the American Political Science Association was misplaced. With the inevitability of generational turnover, cycles of moderation and polarization are likely to be the norm rather than the exception of political dynamics.

To conclude, we return to the motivating question of what causes polarization. Our answer is that it is complicated. We show that the necessary ingredient for polarization is the interaction of voters and elites. Voter updating is necessary, yet on its own doesn't necessarily lead to polarization. It is only when combined with the strategic maneuverings of party elites that a feedback loop is created and polarization occurs. However one interprets responsibility within this relationship, the depth and subtlety of the co-determination in this process points to why researchers have had such difficulty in isolating an individual cause of polarization. Our results establish that a focus on

²⁷The combination of this assumption and the dynamics of party polarization also creates cross-generational patterns that match long-standing empirical findings. The most striking implication is that the propensity to vote strictly increases as citizens age. This is because older generations are, in a sense, captured by the party and pulled with it toward the extreme. A newly born citizen may land at the same moderate location as her mother did, yet the daughter will abstain as the party has polarized and alienated her, whereas the mother updates towards the party position and continues to vote. Notably, this prediction is not dependent on party polarization. All that is required is movement by the parties, and so the same pattern would emerge during a period of party moderation. It is significant, therefore, that higher turnout among older generations is a prominent and persistent feature of the data, dating back to the seminal *Who Votes?* book by [Wolfinger and Rosenstone \(1980\)](#) and continued since then. See <http://www.electproject.org/home/voter-turnout/demographics>.

²⁸Such a cycle of polarization through births and deaths is also consistent with the scattered evidence and popular conception that today it is the old who are the radicals whereas in the 1960's it was the young. This turnover matches exactly the beginning of the modern era of polarization in the 1970's.

the *process* of polarization—and not just on the *state* of polarization—is essential to an understanding of polarization’s origins, its impact, and its future.

References

- ABRAMOWITZ, A. I. (2010): *The Disappearing Center: Engaged Citizens, Polarization, and American Democracy*, Yale University Press.
- ABRAMOWITZ, A. I. AND S. WEBSTER (2016): “The rise of negative partisanship and the nationalization of U.S. elections in the 21st century,” *Electoral Studies*, 41, 12–22.
- ACHARYA, A., M. BLACKWELL, AND M. SEN (2018): “Explaining Preferences from Behavior: A Cognitive Dissonance Approach,” *Journal of Politics*, 80, 400–411.
- ACHEN, C. H. AND L. M. BARTELS (2016): *Democracy for Realists: Why Elections Do Not Produce Responsive Government*, Princeton University Press.
- AKERLOF, G. A. AND W. T. DICKENS (1982): “The Economic Consequences of Cognitive Dissonance,” *The American Economic Review*, 72, 307–319.
- ARONSON, E., C. FRIED, AND J. STONE (1991): “Overcoming Denial and Increasing the Intention to Use Condoms through the Induction of Hypocrisy,” *American Journal of Public Health*, 81, 1636–1638.
- BAGNOLI, M. AND T. BERGSTROM (2005): “Log-Concave Probability and its Applications,” *Economic Theory*, 26, 445–469.
- BEASLEY, R. K. AND M. R. JOSLYN (2001): “Cognitive Dissonance and Post-Decision Attitude Change in Six Presidential Elections,” *Political Psychology*, 22, 521–540.
- BERNHEIM, D., L. BRAGHIERI, A. MARTINEZ-MARQUINA, AND D. ZUCKERMAN (2019): “A Theory of Chosen Preferences,” Working paper.
- BERNHEIM, D. AND A. RANGEL (2009): “Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics,” *The Quarterly Journal of Economics*, 124, 51–104.
- CALVERT, R. L. (1985): “Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence,” *American Journal of Political Science*, 29, 69–95.
- CAMPBELL, A., P. E. CONVERSE, W. E. MILLER, AND D. E. STOKES (1960): *The American Voter*, University of Chicago Press.

- CLARK, T. S. (2009): "Measuring Ideological Polarization on the United States Supreme Court," *Political Research Quarterly*, 62, 146–157.
- CONVERSE, P. E. (1964): *The Nature of Belief Systems in Mass Publics*, New York Press.
- DIMAGGIO, P., J. EVANS, AND B. BRYSON (1996): "Have American's Social Attitudes Become More Polarized?" *American Journal of Sociology*, 102, 690–755.
- ESTEBAN, J.-M. AND D. RAY (1994): "On the Measurement of Polarization," *Econometrica*, 62, 819–851.
- (2012): "Comparing Polarization Measures," in *The Oxford Handbook of the Economics of Peace and Conflict*, ed. by M. R. Garfinkel and S. Skaperdas, Oxford University Press.
- FESTINGER, L. (1962): *A Theory of Cognitive Dissonance*, Stanford University Press.
- FIORINA, M. P., S. J. ABRAMS, AND J. C. POPE (2010): *Culture War? The Myth of a Polarized America*, Longman.
- GENTZKOW, M. (2016): "Polarization in 2016," *Toulouse Network for Information Technology*, white paper.
- GHITZA, Y. AND A. GELMAN (2014): "The Great Society, Reagan's Revolution, and Generations of Presidential Voting," Working paper.
- HALL, A. B. AND D. M. THOMPSON (2018): "Who Punishes Extremist Nominees? Candidate Ideology and Turning Out the Base in US Elections," *American Political Science Review*, 112, 509–524.
- HOTELLING, H. (1929): "Stability in Competition," *Economic Journal*, 39, 41–57.
- JESSEE, S. A. (2009): "Spatial Voting in the 2004 Presidential Election," *American Political Science Review*, 103, 59–81.
- (2010): "Partisan Bias, Political Information and Spatial Voting in the 2008 Presidential Election," *The Journal of Politics*, 72, 327–340.
- KANTROWITZ, R. AND M. M. NEUMANN (2007): "Is the Optimal Rectangle a Square?" *Pi Mu Epsilon Journal*, 12, 405–412.
- KLEMPERER, P. (1987): "Markets with Consumer Switching Costs," *The Quarterly Journal of Economics*, 102, 375–394.
- KÖSZEGI, B. AND M. RABIN (2006): "A Model of Reference-Dependent Preferences," *The Quarterly Journal of Economics*, 121, 1133–1165.

- LENZ, G. S. (2012): *Follow the Leader? How Voters Respond to Politicians' Policies and Performance*, University of Chicago Press.
- LEVENDUSKY, M. (2009): *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*, Chicago Studies in American Politics, Chicago, IL: University of Chicago Press.
- MCCARTY, N. (2019): *Polarization: What Everyone Needs to Know*, Oxford University Press.
- MCDONALD, M. P. AND S. L. POPKIN (2001): "The Myth of the Vanishing Voter," *American Political Science Review*, 95, 963–974.
- MULLAINATHAN, S. AND E. WASHINGTON (2009): "Sticking with Your Vote: Cognitive Dissonance and Political Attitudes," *American Economic Journal: Applied Economics*, 1, 86–111.
- PAGE, S. (2006): "Path Dependence," *Quarterly Journal of Political Science*, 1, 87–115.
- PANAGOPOULOS, C. (2016): "All About that Base: Changing Campaign Strategies in U.S. Presidential Elections," *Party Politics*, 22, 179–190.
- PENN, E. M. (2017): "Inequality, Social Context, and Value Divergence," *The Journal of Politics*, 79, 153–165.
- PIERSON, P. (2000): "Increasing Returns, Path Dependence, and the Study of Politics," *The American Political Science Review*, 94, 251–267.
- (2004): *Politics in Time: History, Institutions, and Social Analysis*, Princeton University Press.
- PUETT, M. AND C. GROSS-LOH (2016): *The Path: What Chinese Philosophers Can Teach Us About the Good Life*, Simon and Schuster.
- ROY, R. AND F. W. J. OLVER (2010): "Elementary Functions," in *NIST Handbook of Mathematical Functions*, ed. by F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, Cambridge, MA: Cambridge University Press, chap. 4.
- SMITHIES, A. (1941): "Optimum Location in Spatial Competition," *The Journal of Political Economy*, 49, 423–439.
- WOLFINGER, R. E. AND S. J. ROSENSTONE (1980): *Who Votes?*, Yale University Press.

APPENDIX

A Equilibrium in the First Period Electoral Competition

For simplicity of notation, we ignore the time subscript to analyze the first period election. Throughout, we assume that policy positions satisfy $d \leq r$. The opposite case never occurs in equilibrium given parties' preferences. These preferences also imply that, in any equilibrium, $D \leq d$ and $r \leq R$. Thus, we treat the *policy space* for both parties as the closed, bounded interval $\mathcal{P} = [D, R]$.

For given $r, d \in \mathcal{P}$, issue voters' support for parties R and D are denoted η^R and η^D , respectively. *Vote total* for each party is

$$V^R(r, d; \lambda, \alpha) = \int_{v \in \eta^R} f(v; \alpha) dv \quad \text{and} \quad V^D(r, d; \lambda, \alpha) = \int_{v \in \eta^D} f(v; \alpha) dv,$$

where $f(v; \alpha)$ is the density function of the logistic distribution with zero mean and scale $\alpha > 0$. Party R 's *vote share* (winning probability) is therefore

$$S^R(r, d; \lambda, \alpha) = \frac{V^R(r, d; \lambda, \alpha)}{V^R(r, d; \lambda, \alpha) + V^D(r, d; \lambda, \alpha)}.$$

Party D 's vote share is of course $1 - S^R(r, d; \lambda, \alpha)$.

Voters' support for each party depends on the value of the tolerance parameter $\lambda > 0$ with respect to the distance $r - d$. There are two different regimes to consider. In *regime A*, the tolerance region λ is sufficiently large relative to the distance $r - d$ so that the parties' intervals of support η^R and η^D have a common boundary; only extreme voters abstain. In *regime B*, tolerance is sufficiently low so that parties' intervals of support do not touch; there is abstention in the middle as well as in the extremes. To be precise:

- *Regime A: abstention only in the extremes.* When $\lambda \geq \frac{r-d}{2}$, we have

$$\eta_A^D = [d - \lambda, \frac{r+d}{2}] \quad \text{and} \quad \eta_A^R = [\frac{r+d}{2}, r + \lambda].$$

- *Regime B: abstention in the middle and in the extremes.* When $\lambda \leq \frac{r-d}{2}$, we have

$$\eta_B^D = [d - \lambda, d + \lambda] \quad \text{and} \quad \eta_B^R = [r - \lambda, r + \lambda].$$

A.1 Vote Share Representation

Given $d \in \mathcal{P}$ and $\lambda, \alpha > 0$, when R chooses a policy position consistent with *regime k* (for $k = A, B$), we denote its vote share by $S_k^R(r, d; \lambda, \alpha)$. To be specific:

- *Regime A:* for $r \in [d, d + 2\lambda] \cap \mathcal{P}$,

$$S_A^R(r, d; \lambda, \alpha) \equiv \frac{F(r + \lambda; \alpha) - F\left(\frac{r+d}{2}; \alpha\right)}{F(r + \lambda; \alpha) - F(d - \lambda; \alpha)}; \quad \text{and}$$

- *Regime B:* for $r \in [d + 2\lambda, +\infty) \cap \mathcal{P}$,

$$S_B^R(r, d; \lambda, \alpha) \equiv \frac{F(r + \lambda; \alpha) - F(r - \lambda; \alpha)}{F(r + \lambda; \alpha) - F(r - \lambda; \alpha) + F(d + \lambda; \alpha) - F(d - \lambda; \alpha)}.$$

It is convenient to rely on the representation of the logistic distribution $F(v; \alpha)$ in terms of the hyperbolic trigonometric functions $\sinh(x)$, $\cosh(x)$ and $\tanh(x)$. See [Appendix C \(online\)](#) for definitions and basic properties of these functions.

Lemma A.1. Fix $d \in \mathcal{P}$ and $\lambda, \alpha > 0$.

- (a) $S_A^R(\cdot, d; \lambda, \alpha)$ can be represented via

$$S_A^R(r, d; \lambda, \alpha) = \left(1 + \frac{\cosh\left(\frac{r+\lambda}{2\alpha}\right)}{\cosh\left(\frac{d-\lambda}{2\alpha}\right)}\right)^{-1}. \quad (5)$$

- (b) $S_B^R(\cdot, d; \lambda, \alpha)$ can be represented via

$$S_B^R(r, d; \lambda, \alpha) = \left(1 + \frac{\cosh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right)}{\cosh\left(\frac{d+\lambda}{2\alpha}\right) \cosh\left(\frac{d-\lambda}{2\alpha}\right)}\right)^{-1}. \quad (6)$$

Proof. Both results follow from tedious manipulations. We relegate details of the proof to [Appendix C \(online\)](#). \square

Using [Lemma A.1](#), it is straightforward to obtain the following.

Lemma A.2. Let $d \in \mathcal{P}$ and $\lambda, \alpha > 0$ be given.

- (a) If $d + 2\lambda \leq R$, then the vote share function $S_B^R(\cdot, d; \lambda, \alpha)$ has a unique maximum r_B^v in $[d + 2\lambda, R]$ given by

$$r_B^v = \begin{cases} 0, & \text{for } 0 < \lambda < \frac{|d|}{2}, \\ d + 2\lambda, & \text{for } \lambda \geq \frac{|d|}{2}. \end{cases}$$

- (b) If $d + 2\lambda > R$, then the domain of $S_B^R(\cdot, d; \lambda, \alpha)$ is empty.

Proof. (a) By assumption the domain of the vote share function for regime B is $[d + 2\lambda, R]$. Using

Equation 6 to differentiate $S_B^R(r, d; \lambda, \alpha)$ with respect to r yields

$$\begin{aligned} \frac{\partial S_B^R(r, d; \lambda, \alpha)}{\partial r} &= -\frac{1}{2\alpha} \left(1 + \frac{\cosh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right)}{\cosh\left(\frac{d+\lambda}{2\alpha}\right) \cosh\left(\frac{d-\lambda}{2\alpha}\right)} \right)^{-2} \\ &\quad \times \frac{\sinh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right) + \cosh\left(\frac{r+\lambda}{2\alpha}\right) \sinh\left(\frac{r-\lambda}{2\alpha}\right)}{\cosh\left(\frac{d+\lambda}{2\alpha}\right) \cosh\left(\frac{d-\lambda}{2\alpha}\right)} \\ &= -\frac{1}{2\alpha} \left[S_B^R(r, d; \lambda, \alpha) \right]^2 \frac{\sinh\left(\frac{r}{\alpha}\right)}{\cosh\left(\frac{d+\lambda}{2\alpha}\right) \cosh\left(\frac{d-\lambda}{2\alpha}\right)}. \end{aligned} \quad (7)$$

The last equality in the above expression uses the property (h) in [Theorem C.1](#) in [Appendix C \(online\)](#). Note $\cosh(x) \geq 1$ for all x , and further $\sinh(x) < 0$ for all $x < 0$, $\sinh(x) > 0$ for all $x > 0$, and $\sinh(0) = 0$. Thus, from the expression for the partial derivative of $S_B^R(r, d; \lambda)$ with respect to r , one sees that

$$\begin{aligned} \frac{\partial S_B^R(r, d; \lambda)}{\partial r} &> 0, & \text{for } r < 0, \\ &< 0, & \text{for } r > 0, \\ &= 0, & \text{for } r = 0. \end{aligned}$$

Since $S_B^R(r, d; \lambda)$ is defined for $r \geq d + 2\lambda$, one has that for $d + 2\lambda < 0$ the maximizer is 0 and for $d + 2\lambda \geq 0$ the maximizer is $d + 2\lambda$.

(b) Immediate from the fact that $d + 2\lambda > R$. □

An immediate implication of [Lemma A.2](#) is that $S_B^R(\cdot, d; \lambda, \alpha)$ is maximized at $r_B^v \geq 0$.

A.2 Expected Utility as a Product Function

With linear preferences, R 's expected utility given $d \leq r$ can be written as

$$\mathbb{E}U^R(r, d; \lambda, \alpha, \beta) = S^R(r, d; \lambda, \alpha) (r - d + \beta) - (R - d). \quad (8)$$

Thus, R 's expected utility is proportional to the difference between policy positions. Increasing r yields a higher payoff, conditional on winning, but can negatively affect R 's winning probability. Since the expected utility is a product function, in our analysis we employ the following result—for a proof see [Kantrowitz and Neumann \(2007\)](#).

Theorem A.1. *Consider log-concave continuous functions g_1, g_2 defined on a closed bounded interval $[a, b]$, and suppose that $g_i(x) > 0$ for all $x \in (a, b)$, for $i = 1, 2$. If one of these functions is strictly log-concave, then there exists a point $x^* \in [a, b]$ such that the product function $h = g_1 g_2$ is strictly increasing on $[a, x^*]$ and strictly decreasing on $[x^*, b]$.*

Clearly, an affine function is strictly log-concave. Thus, to use [Theorem A.1](#) suffices to show that S_k^R is log-concave on r , for $k = A, B$. The density $f(\cdot; \alpha)$ of the logistic distribution is log-concave ([Bagnoli and Bergstrom, 2005](#)). Thus, it follows that the distribution $F(\cdot; \alpha)$ is log-concave. Unfortunately these facts do not translate immediately into the log-concavity of the vote share functions for regimes A and B . Nonetheless, we are able to derive the following.

Lemma A.3. *Fix $d \in \mathcal{P}$ and $\lambda, \alpha > 0$.*

- (a) $S_A^R(\cdot, d; \lambda, \alpha)$ is strictly log-concave.
- (b) $S_B^R(\cdot, d; \lambda, \alpha)$ is strictly log-concave.

Proof. One has to verify that, for each $k = A, B$, the function $\log S_k^R(r, d; \lambda, \alpha)$ is concave on r in its respective domain. We do so by checking the sign of the second derivatives. Details can be found in [Appendix C \(online\)](#). \square

A.3 Equilibrium

Obtaining Nash equilibria in the first period electoral competition game is complicated by the fact that there can be an endogenous switch between regimes A and B , depending on what policy positions parties announce. We proceed by treating these regimes as two independent games. That is, we fix λ, α and β , and restrict the policy space for each party to correspond to either regime A or regime B . Focusing on party R , we obtain R 's best response function in each regime and show that this leads to a unique symmetric equilibrium for each case. Finally, we put both regimes together and construct the unique equilibrium for the true electoral competition game between R and D in period one.

Ignoring the constant $-(R - d)$, party R chooses r in regime $k = A, B$ to maximize

$$\mathbb{E}U_k^R(\cdot, d; \lambda, \alpha, \beta) = S_k^R(r, d; \lambda, \alpha)(r - d + \beta)$$

on its respective domain.

Proposition A.1. *Let $\lambda, \alpha > 0$, $\beta \geq 0$, and $d \in \mathcal{P}$ be given.*

- (a) *The expected utility function $\mathbb{E}U_A^R(\cdot, d; \lambda, \alpha, \beta)$ defined for r on the interval $[d, d + 2\lambda] \cap \mathcal{P}$ has a unique maximum*

$$r_A^*(d; \lambda, \alpha, \beta).$$

Moreover, $\mathbb{E}U_A^R(\cdot, d; \lambda, \alpha, \beta)$ is strictly increasing to the left of $r_A^(d; \lambda, \alpha, \beta)$ and strictly decreasing to the right of $r_A^*(d; \lambda, \alpha, \beta)$.*

- (b) *The expected utility function $\mathbb{E}U_B^R(\cdot, d; \lambda, \alpha, \beta)$ defined for r on the interval $[d + 2\lambda, +\infty] \cap \mathcal{P}$ has a unique maximum*

$$r_B^*(d; \lambda, \alpha, \beta).$$

Moreover, $\mathbb{E}U_B^R(\cdot, d; \lambda, \alpha, \beta)$ is strictly increasing to the left of $r_B^(d; \lambda, \alpha, \beta)$ and strictly decreasing to the right of $r_B^*(d; \lambda, \alpha, \beta)$.*

Proof. Both parts (a) and (b) follow immediately from [Theorem A.1](#). \square

An important implication of [Proposition A.1](#) is that first order conditions are necessary as well as sufficient to obtain an interior maximizer of $\mathbb{E}U_k^R(\cdot, d; \lambda, \alpha, \beta)$.

A.4 Finding Nash Equilibria for Regime A

The optimal policy position $r_A^*(d; \lambda, \alpha, \beta)$ is never at the lower boundary of its domain, when the value of holding office is lower than twice the scale of the logistic distribution.

Lemma A.4. *For all $\lambda, \alpha > 0$, $0 \leq \beta < 2\alpha$ and $d \in \mathcal{P}$, one has that $r_A^*(d; \lambda, \alpha, \beta) > d$.*

Proof. First note that using [Equation 5](#) to differentiate $S_A^R(r, d; \lambda, \alpha)$ with respect to r yields

$$\begin{aligned} \frac{\partial S_A^R(r, d; \lambda, \alpha)}{\partial r} &= -\frac{1}{2\alpha} \left(1 + \frac{\cosh\left(\frac{r+\lambda}{2\alpha}\right)}{\cosh\left(\frac{d-\lambda}{2\alpha}\right)} \right)^{-2} \frac{\sinh\left(\frac{r+\lambda}{2\alpha}\right)}{\cosh\left(\frac{d-\lambda}{2\alpha}\right)} \\ &= -\frac{1}{2\alpha} \left[S_A^R(r, d; \lambda, \alpha) \right]^2 \frac{\sinh\left(\frac{r+\lambda}{2\alpha}\right)}{\cosh\left(\frac{d-\lambda}{2\alpha}\right)}. \end{aligned} \quad (9)$$

Using [Equation 9](#) and partially differentiating $\mathbb{E}U_A^R$ with respect to r (allowing for directional derivatives on the boundaries) obtains:

$$\begin{aligned} \frac{\partial \mathbb{E}U_A^R(r, d; \lambda, \alpha, \beta)}{\partial r} &= S_A^R(r, d; \lambda, \alpha) - \frac{r-d+\beta}{2\alpha} \left[S_A^R(r, d; \lambda, \alpha) \right]^2 \frac{\sinh\left(\frac{r+\lambda}{2\alpha}\right)}{\cosh\left(\frac{d-\lambda}{2\alpha}\right)} \\ &= S_A^R(r, d; \lambda, \alpha) \left\{ 1 - \frac{r-d+\beta}{2\alpha} S_A^R(r, d; \lambda, \alpha) \frac{\sinh\left(\frac{r+\lambda}{2\alpha}\right)}{\cosh\left(\frac{d-\lambda}{2\alpha}\right)} \right\} \\ &= S_A^R(r, d; \lambda, \alpha) \left\{ 1 - \frac{r-d+\beta}{2\alpha} \frac{\sinh\left(\frac{r+\lambda}{2\alpha}\right)}{\cosh\left(\frac{r+\lambda}{2\alpha}\right) + \cosh\left(\frac{d-\lambda}{2\alpha}\right)} \right\}. \end{aligned}$$

Immediately from evaluating this expression at $r = d$, we obtain

$$\left. \frac{\partial \mathbb{E}U_A^R(r, d; \lambda, \alpha, \beta)}{\partial r} \right|_{r=d} = S_A^R(d, d; \lambda, \alpha) \left\{ 1 - \frac{\beta}{2\alpha} \frac{\sinh\left(\frac{d+\lambda}{2\alpha}\right)}{\cosh\left(\frac{d+\lambda}{2\alpha}\right) + \cosh\left(\frac{d-\lambda}{2\alpha}\right)} \right\}.$$

The vote share $S_A^R(d, d; \lambda, \alpha)$ is strictly positive. Using identities (g) and (h) of [Theorem C.1](#) in

Appendix C (online), write the expression in the brackets of the above equation as

$$\begin{aligned}
1 - \frac{\beta}{2\alpha} \frac{\sinh\left(\frac{d+\lambda}{2\alpha}\right)}{\cosh\left(\frac{d+\lambda}{2\alpha}\right) + \cosh\left(\frac{d-\lambda}{2\alpha}\right)} &= 1 - \frac{\beta}{4\alpha} \frac{\sinh\left(\frac{d}{2\alpha}\right) \cosh\left(\frac{\lambda}{2\alpha}\right) + \cosh\left(\frac{d}{2\alpha}\right) \sinh\left(\frac{\lambda}{2\alpha}\right)}{\cosh\left(\frac{d}{2\alpha}\right) \cosh\left(\frac{\lambda}{2\alpha}\right)} \\
&= 1 - \frac{\beta}{4\alpha} \left(\frac{\sinh\left(\frac{d}{2\alpha}\right)}{\cosh\left(\frac{d}{2\alpha}\right)} + \frac{\sinh\left(\frac{\lambda}{2\alpha}\right)}{\cosh\left(\frac{\lambda}{2\alpha}\right)} \right) \\
&= 1 - \frac{\beta}{4\alpha} \left(\tanh\left(\frac{d}{2\alpha}\right) + \tanh\left(\frac{\lambda}{2\alpha}\right) \right).
\end{aligned}$$

Since $1 > \tanh(x) > -1$, the last expression is greater than $1 - \beta/2\alpha$. Thus, the marginal utility is strictly positive at $r = d$ as long as $2\alpha > \beta$. Hence $r_A^*(d; \lambda, \alpha, \beta) > d$. \square

We focus on the case where $r \in (d, d + 2\lambda] \cap \mathcal{P}$. In an interior solution, the first order conditions can be expressed as

$$\mathcal{G}(r, d; \lambda, \alpha, \beta) \equiv 1 - \frac{r - d + \beta}{2\alpha} \frac{\sinh\left(\frac{r+\lambda}{2\alpha}\right)}{\cosh\left(\frac{r+\lambda}{2\alpha}\right) + \cosh\left(\frac{d-\lambda}{2\alpha}\right)} = 0. \quad (10)$$

Let $r_A^\circ(d; \lambda, \alpha, \beta)$ denote the interior maximizer; i.e., the implicit solution to Equation 10. This is illustrated in Figure 9, where the horizontal axis corresponds to values of d and the vertical axis to values of $r_A^\circ(d; \lambda, \alpha, \beta)$, for $\lambda = 2.5$, $\alpha = 1$ and $\beta = 0$. We now express R 's best response function in regime A as

$$r_A^*(d; \lambda, \alpha, \beta) = \min \{r_A^\circ(d; \lambda, \alpha, \beta), d + 2\lambda, R\}. \quad (11)$$

We look for symmetric equilibria. Assume the symmetric equilibrium is the corner solution, i.e., $\min\{d + 2\lambda, R\}$. Then one has either $r = d + 2\lambda = -r + 2\lambda$, and thus in equilibrium $r_A^*(\lambda, \alpha, \beta) = \lambda$, or $r_A^*(\lambda, \alpha, \beta) = R$. Assume now that the symmetric equilibrium is interior. Then substitute $d = -r$ in Equation 10 to obtain

$$1 = \frac{2r + \beta}{2\alpha} \frac{\sinh\left(\frac{r+\lambda}{2}\right)}{\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(-\frac{r+\lambda}{2}\right)} = \frac{2r + \beta}{4\alpha} \frac{\sinh\left(\frac{r+\lambda}{2}\right)}{\cosh\left(\frac{r+\lambda}{2}\right)},$$

where the last equality follows from the fact that the hyperbolic cosine function is even. Therefore, the interior symmetric equilibrium strategy is characterized by the equation

$$1 = \left(\frac{r}{2\alpha} + \frac{\beta}{4\alpha} \right) \tanh\left(\frac{r + \lambda}{2\alpha}\right). \quad (12)$$

This leads to the following proposition.

Proposition A.2. *Given $\lambda, \alpha > 0$ and $0 \leq \beta < 2\alpha$, the symmetric equilibrium in regime A is*

$$r_A^*(\lambda, \alpha, \beta) = \min \{r_A^\circ(\lambda, \alpha, \beta), \lambda, R\} \quad \text{and} \quad d_A^*(\lambda, \alpha, \beta) = -r_A^*(\lambda, \alpha, \beta),$$

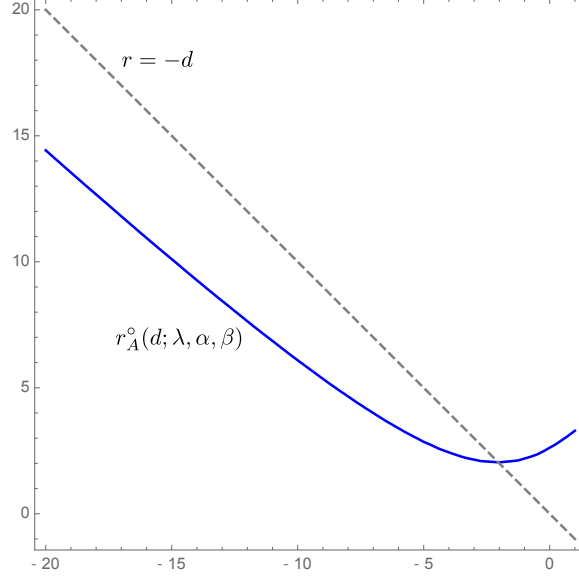


Figure 9: $r_A^o(d; \lambda, \alpha, \beta)$ for $\lambda = 2.5, \alpha = 1, \beta = 0$. The horizontal axis represents values of d , the vertical axis represents values of r .

where $r_A^o(\lambda, \alpha, \beta)$ is implicitly defined by [Equation 12](#). Moreover, this equilibrium is unique.

Proof. Our previous analysis establishes the existence and characterization of the symmetric equilibrium. We briefly mention here the arguments required to show uniqueness, deferring the details to [Appendix C \(online\)](#).

Because parties have opposite symmetric preferences and the logistic distribution is symmetric around its mean, party D 's best response function is the negative of R 's best response function. Thus, it suffices to show that R 's best response function in [Equation 11](#), as a function of d , is a contraction. By the Contraction Mapping theorem, this implies that the equilibrium is unique. \square

We can now fully describe the unique symmetric equilibrium in regime A using some comparative statics. [Lemma C.2](#) in [Appendix C](#) applies the Implicit Function theorem to [Equation 10](#) to obtain

$$\frac{\partial r_A^o(\lambda, \alpha, \beta)}{\partial \lambda} < 0.$$

Moreover, for $\lambda \rightarrow 0$ [Equation 12](#) shows that $r_A^o > 0$. Otherwise, if $r_A^o = 0$ we obtain $1 = 0$ because $\sinh(0) = 0$ and $\cosh(0) = 1$.²⁹ The symmetric equilibrium policy starts at $r_A^*(\lambda, \alpha, \beta) = \lambda$ for small values of λ and is strictly increasing. It changes from the corner solution to the interior at a tolerance level λ^* such that $\lambda^* = r_A^o(\lambda^*, \alpha, \beta)$, as long as $R > \lambda^*$, otherwise it remains fixed at R .³⁰ For values of λ above λ^* , the symmetric equilibrium in regime A is $r_A^*(\lambda, \alpha, \beta) = r_A^o(\lambda, \alpha, \beta)$ and is strictly decreasing, although asymptotic. Indeed, since

²⁹One can solve for r_A^o numerically, given values of α and β . For example, when $\alpha = 1$ and $\beta = 0$, one obtains $r_A^o(0, 1, 0) \approx 2.399357280$.

³⁰Using [Equation 12](#) again, one can numerically estimate the value of λ^* , for $\alpha = 1$ and $\beta = 0$, to be $\lambda^* \approx 2.065338138$.

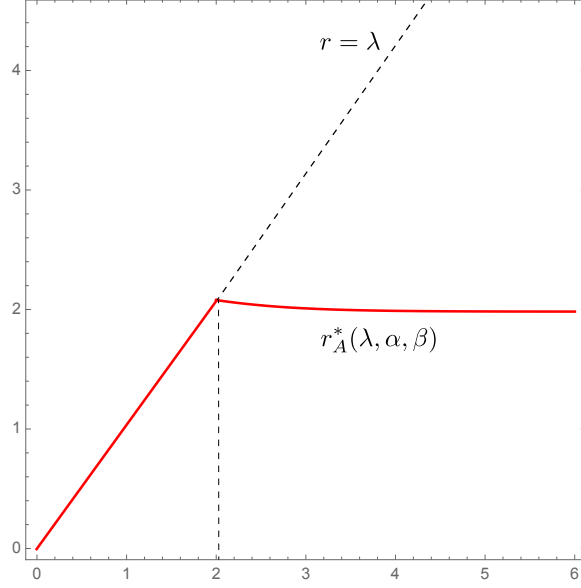


Figure 10: Symmetric equilibrium for Regime A, for $\alpha = 1, \beta = 0$. The horizontal axis represents λ , the vertical axis represents r .

$\frac{\sinh(x)}{\cosh(x)} = \lim_{x \rightarrow \infty} \tanh(x) = 1$, one sees again from Equation 12 that $r^*(\lambda, \alpha, \beta) \rightarrow 2\alpha - \beta/2$ as $\lambda \rightarrow \infty$. This is illustrated in Figure 10, which plots the value of the tolerance parameter in the horizontal axis and the equilibrium strategy r_A^* in the vertical axis, for $\alpha = 1$ and $\beta = 0$.

A.5 Finding NE for Regime B

Fix $\lambda, \alpha > 0$, $\beta \geq 0$ and $d \in \mathcal{P}$. Recall that R 's policy domain in regime B is the interval $[d + 2\lambda, +\infty) \cap \mathcal{P}$. Assume that $d + 2\lambda \leq R$, else the domain of regime B is empty. Focusing first on the interior solution, we use Equation 7 and differentiate $\mathbb{E}U_B^R(r, d; \lambda, \alpha, \beta)$ with respect to r

(allowing for directional derivatives on the boundaries). This obtains:

$$\begin{aligned}
\frac{\partial \mathbb{E}U_B^R(r, d; \lambda, \alpha, \beta)}{\partial r} &= S_B^R(r, d; \lambda, \alpha) - \frac{r - d + \beta}{2\alpha} \left[S_B^R(r, d; \lambda, \alpha) \right]^2 \frac{\sinh\left(\frac{r}{\alpha}\right)}{\cosh\left(\frac{d+\lambda}{2\alpha}\right) \cosh\left(\frac{d-\lambda}{2\alpha}\right)} \\
&= S_B^R(r, d; \lambda, \alpha) \\
&\quad \times \left\{ 1 - \frac{r - d + \beta}{2\alpha} \left(1 + \frac{\cosh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right)}{\cosh\left(\frac{d+\lambda}{2\alpha}\right) \cosh\left(\frac{d-\lambda}{2\alpha}\right)} \right)^{-1} \frac{\sinh\left(\frac{r}{\alpha}\right)}{\cosh\left(\frac{d+\lambda}{2\alpha}\right) \cosh\left(\frac{d-\lambda}{2\alpha}\right)} \right\} \\
&= S_B^R(r, d; \lambda, \alpha) \\
&\quad \times \left\{ 1 - \frac{r - d + \beta}{2\alpha} \frac{\sinh\left(\frac{r}{\alpha}\right)}{\cosh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right) + \cosh\left(\frac{d+\lambda}{2\alpha}\right) \cosh\left(\frac{d-\lambda}{2\alpha}\right)} \right\} \\
&= S_B^R(r, d; \lambda, \alpha) \left\{ 1 - \frac{r - d + \beta}{\alpha} \frac{\sinh\left(\frac{r}{\alpha}\right)}{\cosh\left(\frac{r}{\alpha}\right) + \cosh\left(\frac{d}{\alpha}\right) + 2 \cosh\left(\frac{\lambda}{\alpha}\right)} \right\},
\end{aligned}$$

where the last equality uses identity (g) from [Theorem C.1](#) in [Appendix C](#) (online).

In an interior solution, because vote share is strictly positive, the FOC for regime B can be expressed as:

$$\mathcal{H}(r, d; \lambda, \alpha, \beta) \equiv 1 - \frac{r - d + \beta}{\alpha} \frac{\sinh\left(\frac{r}{\alpha}\right)}{\cosh\left(\frac{r}{\alpha}\right) + \cosh\left(\frac{d}{\alpha}\right) + 2 \cosh\left(\frac{\lambda}{\alpha}\right)} = 0. \quad (13)$$

Let $r_B^\circ(d; \lambda, \alpha, \beta)$ denote the interior maximizer; i.e., the solution to [Equation 13](#) above. This is illustrated in [Figure 11](#), where the horizontal axis corresponds to values of d and the vertical axis to values of $r_B^\circ(d; \lambda, \alpha, \beta)$ for $\lambda = 1.5, \alpha = 1, \beta = 0$.

In regime B we cannot rule out a priori either corner solution. To spare on notation, let $\text{mid}\{a, b, c\}$ denote the middle value of three real numbers a, b, c . Express now R 's best response function for regime B as

$$r_B^*(d; \lambda, \alpha, \beta) = \text{mid}\{r_B^\circ(d; \lambda, \alpha, \beta), d + 2\lambda, R\}. \quad (14)$$

We look again for the symmetric Nash equilibria. It is immediate to see that a corner solution yields to $r_B^*(\lambda, \alpha, \beta) = \lambda$ or $r_B^*(\lambda, \alpha, \beta) = R$ as a symmetric equilibrium. Assume instead that the symmetric equilibrium is interior. Substituting $d = -r$ in [Equation 13](#) and using the fact that the hyperbolic cosine function is even, obtains

$$1 = \left(\frac{r}{\alpha} + \frac{\beta}{2\alpha} \right) \frac{\sinh\left(\frac{r}{\alpha}\right)}{\cosh\left(\frac{r}{\alpha}\right) + \cosh\left(\frac{\lambda}{\alpha}\right)}.$$

Replacing $\frac{r}{\alpha} = \frac{r+\lambda}{2\alpha} + \frac{r-\lambda}{2\alpha}$ in the numerator, and applying identities (h) and (g) of [Theorem C.1](#) in

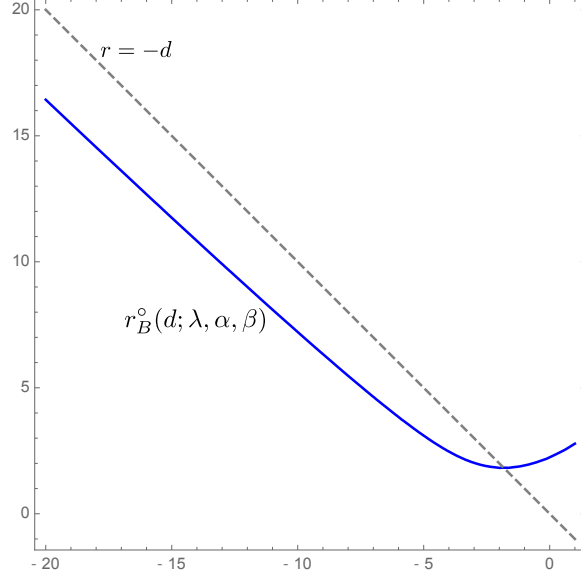


Figure 11: $r_B^o(d; \lambda, \alpha, \beta)$, for $\lambda = 1.5, \alpha = 1, \beta = 0$. The horizontal axis represents values of d , the vertical axis represents values of r

Appendix C in the numerator and denominator, respectively, obtains

$$1 = \left(\frac{r}{2\alpha} + \frac{\beta}{4\alpha} \right) \left\{ \tanh \left(\frac{r + \lambda}{2\alpha} \right) + \tanh \left(\frac{r - \lambda}{2\alpha} \right) \right\}. \quad (15)$$

This equation characterizes the interior equilibrium. We obtain the following.

Proposition A.3. *Given $\lambda, \alpha > 0$ and $\beta \geq 0$, the symmetric equilibrium in regime B is given by*

$$r_B^*(\lambda, \alpha, \beta) = \text{mid} \{ r_B^o(\lambda, \alpha, \beta), \lambda, R \} \quad \text{and} \quad d_B^*(\lambda, \alpha, \beta) = -r_B^*(\lambda, \alpha, \beta),$$

where $r_B^o(\lambda, \alpha, \beta)$ is implicitly defined by Equation 15. Moreover, this equilibrium is unique.

Proof. We have already argued existence and characterization of a symmetric equilibrium. The arguments for uniqueness are similar to those of Proposition A.2. We defer the details to Appendix C (online). \square

As in the previous case, we can fully describe the symmetric equilibrium in regime B. We show in Lemma C.2 in the Appendix C that the Implicit Function theorem on Equation 15 obtains

$$\frac{\partial r_B^o(\lambda, \alpha, \beta)}{\partial \lambda} > 0.$$

Moreover, for $\lambda \rightarrow 0$, Equation 15 shows that $r_B^o > 0$.³¹ The symmetric equilibrium policy in regime B starts at $r_B^*(\lambda, \alpha, \beta) = r_B^o(\lambda, \alpha, \beta)$ for small values of λ and is strictly increasing. It changes from the interior solution to the corner at the tolerance level λ^* such that

³¹One can also solve for r_B^o numerically, given values of α and β . When $\alpha = 1$ and $\beta = 0$, one obtains $r_B^o(0, 1, 0) \approx 1.543404638$.

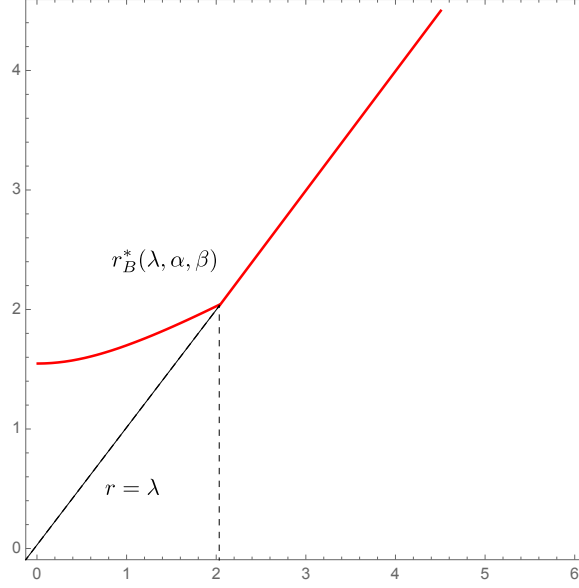


Figure 12: Symmetric equilibrium for Regime B, for $\alpha = 1, \beta = 0$. The horizontal axis represents λ , the vertical axis represents r .

$\lambda^* = r_B^\circ(\lambda^*, \alpha, \beta) = r_A^\circ(\lambda^*, \alpha, \beta)$. This last equality follows from Equation 15 together with Equation 12. The symmetric equilibrium is then $r_B^*(\lambda, \alpha, \beta) = \lambda$ for $\lambda \leq R$, and $r_B^*(\lambda, \alpha, \beta) = R$ for $\lambda > R$. This equilibrium is illustrated in Figure 12, for $\alpha = 1$ and $\beta = 0$ (for a large enough value of R).

A.6 Nash Equilibrium of the First Period Election

Fix a tolerance level $\lambda > 0$, and parameters $\alpha > 0, \beta \geq 0$. To find the Nash equilibria of the first period electoral competition game, we consider candidate R 's actual expected utility, which is given by

$$\mathbb{E}U^R(\cdot, d; \lambda, \alpha, \beta) = \begin{cases} \mathbb{E}U_A^R(\cdot, d; \lambda, \alpha, \beta) & \text{if } d \leq r \leq d + 2\lambda, \\ \mathbb{E}U_B^R(\cdot, d; \lambda, \alpha, \beta) & \text{if } d + 2\lambda \leq r \leq R. \end{cases}$$

If there were no regime switch at all, the NE would be given by either the symmetric solution for regime A or for regime B. Yet there may be some regime switch, so that for some specific value of λ , R 's best response switches from $r_A^*(d; \lambda, \alpha, \beta)$ to $r_B^*(d; \lambda, \alpha, \beta)$ at some given d , or vice versa. In that case, the best response function for the actual electoral competition game may be discontinuous, so one would have to consider mixed strategies to ensure existence of Nash equilibria in the first period electoral competition game. We show below that a unique, symmetric, equilibrium does exist. To rule out uninteresting corner solutions, i.e., where $r = 0$ or $r = R$, we impose the following assumption.

Assumption A.1. Let $\alpha, \lambda > 0$ and $0 \leq \beta < 2\alpha$. Further, let $-D = R > \lambda^*$, where λ^* is the tolerance level that solves

$$\tanh\left(\frac{\lambda^*}{\alpha}\right) = \frac{4\alpha}{2\lambda^* + \beta}.$$

Remark. Note that λ^* defined above is such that $\lambda^* = r_B^\circ(\lambda^*, \alpha, \beta) = r_A^\circ(\lambda^*, \alpha, \beta)$ —see Eq. (15) and Eq. (12). This is also the value of the tolerance parameter given in Equation 2.

The following proposition, which corresponds to Proposition 1 in the main text, fully describes the equilibrium of the electoral competition game in period 1.

Proposition A.4. *Under Assumption A.1, the unique equilibrium of the first period electoral competition game is the strategy profile $(r^*(\lambda, \alpha, \beta), d^*(\lambda, \alpha, \beta))$ such that*

$$r^*(\lambda, \alpha, \beta) = \begin{cases} r_B^\circ(\lambda, \alpha, \beta) & : 0 < \lambda < \lambda^*, \\ \lambda^* & : \lambda = \lambda^*, \\ r_A^\circ(\lambda, \alpha, \beta) & : \lambda > \lambda^*, \end{cases} \quad \text{and} \quad d^*(\lambda, \alpha, \beta) = -r^*(\lambda, \alpha, \beta),$$

where $0 < r_B^\circ(\lambda, \alpha, \beta) < R$ is implicitly defined by Equation 15 and $0 < r_A^\circ(\lambda, \alpha, \beta) < R$ is implicitly defined by Equation 12.

Proof. Due to Assumption A.1, we ignore $r = 0$ or $r = R$ in the proof of the proposition. Building on Proposition A.2 and Proposition A.3, we divide the analysis into four exhaustive cases and exploit the fact that $\mathbb{E}U_k^R(\cdot, d; \lambda, \alpha, \beta)$ is strictly increasing to the left of $r_k^*(d; \lambda, \alpha, \beta)$ and strictly decreasing to the right of $r_k^*(d; \lambda, \alpha, \beta)$. Fix $d \in [D, R]$, $\lambda, \alpha > 0$, and $0 \leq \beta < 2\alpha$.

Case 1—both regimes have a corner solution:

$$r_B^\circ(d; \lambda, \alpha, \beta) < r_B^*(d; \lambda, \alpha, \beta) = d + 2\lambda = r_A^*(d; \lambda, \alpha, \beta) < r_A^\circ(d; \lambda, \alpha, \beta).$$

It follows that $\mathbb{E}U_A^R$ is strictly increasing on its domain and $\mathbb{E}U_B^R$ is strictly decreasing on its domain. Therefore, the actual expected utility $\mathbb{E}U^R$ has a single peak at $r = d + 2\lambda$. Party R's unique best response is therefore $r^*(d; \lambda, \alpha, \beta) = d + 2\lambda$.

Case 2—regime A has a corner solution and regime B an interior solution:

$$r_A^*(d; \lambda, \alpha, \beta) = d + 2\lambda < r_B^\circ(d; \lambda, \alpha, \beta) = r_B^*(d; \lambda, \alpha, \beta).$$

Note also that $d + 2\lambda < r_A^\circ(d; \lambda, \alpha, \beta)$. It follows that $\mathbb{E}U_A^R$ is strictly increasing on its domain and $\mathbb{E}U_B^R$ is strictly increasing around $d + 2\lambda$ and has a single peak at $r_B^\circ(d; \lambda, \alpha, \beta)$. Therefore, the actual expected utility $\mathbb{E}U^R$ has a single peak at $r_B^\circ(d; \lambda, \alpha, \beta)$. Party R's unique best response is therefore $r^*(d; \lambda, \alpha, \beta) = r_B^\circ(d; \lambda, \alpha, \beta)$.

Case 3—regime A has an interior solution and regime B a corner solution:

$$r_A^*(d; \lambda, \alpha, \beta) = r_A^\circ(d; \lambda, \alpha, \beta) < d + 2\lambda = r_B^*(d; \lambda, \alpha, \beta).$$

We must also have $r_B^\circ(d; \lambda, \alpha, \beta) < d + 2\lambda$. It follows that $\mathbb{E}U_A^R$ has a single peak at $r_A^\circ(d; \lambda, \alpha, \beta)$ but $\mathbb{E}U_B^R$ is strictly decreasing on its domain. The actual expected utility $\mathbb{E}U^R$ has a single peak at $r_A^\circ(d; \lambda, \alpha, \beta)$. Therefore R's unique best response is $r^*(d; \lambda, \alpha, \beta) = r_A^\circ(d; \lambda, \alpha, \beta)$.

Case 4—both regimes have an interior solution:

$$r_A^*(d; \lambda, \alpha, \beta) = r_A^\circ(d; \lambda, \alpha, \beta) < d + 2\lambda < r_B^\circ(d; \lambda, \alpha, \beta) = r_B^*(d; \lambda, \alpha, \beta).$$

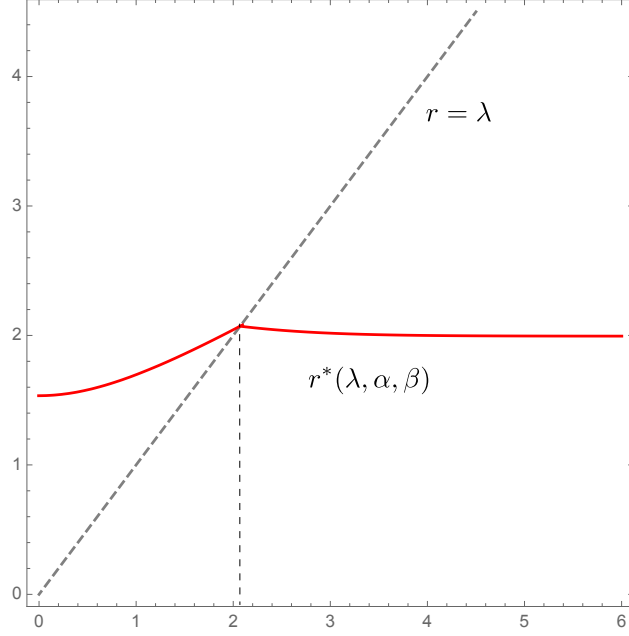


Figure 13: Equilibrium for the Electoral Competition Game, for $\alpha = 1, \beta = 0$. The horizontal axis represents λ , the vertical axis represents r .

Both $\mathbb{E}U_A^R$ and $\mathbb{E}U_B^R$ have a single peak in the interior of their respective domains, and the actual expected utility $\mathbb{E}U^R$ has two (local) maxima. Note however that this case is incompatible with a symmetric equilibrium. Indeed, because in each of the $\mathbb{E}U_k^R$ functions the maximum is interior, it must be the case that in the symmetric equilibrium $-d = r_A^\circ(d; \lambda, \alpha, \beta) < r_B^\circ(d; \lambda, \alpha, \beta) < -d$. Thus, this case will never occur in equilibrium.

We can now describe the symmetric equilibrium of the first period electoral competition game as a function of λ , for fixed parameter values $\alpha > 0$ and $0 \leq \beta < 2\alpha$. For $\lambda < \lambda^*$, Case 2 is relevant and we have $r^*(d; \lambda, \alpha, \beta) = r_B^\circ(d; \lambda, \alpha, \beta)$. Thus, the symmetric equilibrium strategy $r^*(\lambda, \alpha, \beta)$ is characterized by Equation 15. For $\lambda > \lambda^*$, Case 3 is relevant and we have $r^*(d; \lambda, \alpha, \beta) = r_A^\circ(d; \lambda, \alpha, \beta)$. Thus, the symmetric equilibrium strategy $r^*(\lambda, \alpha, \beta)$ is characterized by Equation 12. For $\lambda = \lambda^*$, Case 1 is relevant and thus the symmetric equilibrium yields $r^*(\lambda^*, \alpha, \beta) = -d^*(\lambda^*, \alpha, \beta) = \lambda^*$, where λ^* is given by either Equation 12 or Equation 15 evaluated at $r = \lambda$. \square

The unique equilibrium of the first period electoral game is illustrated in Figure 13, where we plot $r^*(\lambda, \alpha, \beta)$ as a function of λ , for parameter values $\alpha = 0$ and $\beta = 1$.

A.7 Comparative Statics

Higher value for holding office Recall β , the value derived from being in office, is assumed to be $0 \leq \beta < 2\alpha$.³² The first observation is that the effect of a higher β is a downward shift of R 's best response function in both regime A and B . A higher value for office thus implies more convergence to the center and also, albeit less importantly, the fact that the switch between regimes A and B occurs closer to the median. In other words, the cutoff value of λ^* where the equilibrium changes from regime B to regime A is also decreasing in β .

What about different values of β ? Suppose for instance that $\beta^R > 0$ and $\beta^D = 0$. Then the shift in the FOC occurs only for party R . This will lead to asymmetric equilibria. Even though one has to be careful in considering the parameter values for which asymmetric equilibria would exist, the nature of the results should not change.

Higher scale in the logistic distributions In both regime A and B , a larger value of the scale parameter α shifts the FOC curves upwards. This immediately implies that a higher scale in the logistic distribution pushes the symmetric equilibrium to the extremes. Intuitively, a larger scale means a larger proportion of issue voters can be found in the flanks, and thus parties have stronger incentives to move out of the center.

We summarize these findings in the following corollary.

Corollary A.1. *Let $\lambda, \alpha > 0$ and $0 \leq \beta < 2\alpha$.*

- (a) *The first period equilibrium position $r^*(\lambda, \alpha, \beta)$ is decreasing in β .*
- (b) *The first period equilibrium position $r^*(\lambda, \alpha, \beta)$ is increasing in α .*

Proof. (a) Directly from Equation 12 and Equation 15 one sees that the solution to each of these expressions is decreasing in β . Note also that the switch from regime B to regime A occurs at λ^* that solves Equation 2. Clearly, an immediate application of the Implicit Function Theorem shows that the value of λ^* is decreasing in β .

(b) Directly from Equation 12 and Equation 15 one sees that the solution to each of these expressions is increasing in α . □

B Equilibrium in the Second and Subsequent Elections

Here we gather the proofs of the results for the second and subsequent elections. For reasons of space, we omit proofs of results that are immediate.

³²With values of β larger than 2α , the motivation to gain office could be sufficiently high to generate an equilibrium with $r = d = 0$.

Proof of Proposition 4. Suppose that $\lambda \leq \lambda^*$. Recall that in this case the equilibrium in the first period election consists of policy positions $r_1^* \in (0, R)$ and $d_1^* = -r_1^*$ such that there is abstention in the middle and in the extremes (cf. Proposition A.4). We implicitly assume that party R 's ideal point is sufficiently large and show that its equilibrium position in the second election is $r_2^* = r_1^* + \lambda\tau$. This follows from three key observations.

The first observation is that, because the end points of the interval of support in the first election are $r_1^* - \lambda$ and $r_1^* + \lambda$ and the updating rule takes the form given in Eq. (1), any position that is more than $\lambda\tau$ to the left or to the right of r_1^* leaves party R with the same vote total it would have had with said position in the first election. Thus, there is no equilibrium in which both parties choose positions further than $\lambda\tau$ to the left or to the right of their first period equilibrium positions. Suppose, then, that D locates at d_2 in the second election, a position that is at most $\lambda\tau$ left or right of d_1^* .

The second key observation is that R 's first period best response function, which in this case corresponds to what we have referred to as Regime B , attains a unique minimum at r_1^* . Thus, for any location $d \neq d_1^*$ chosen by D in period 1, party R 's best response is some $\hat{r} > r_1^*$. We formally state and prove this observation as Lemma C.1 in Appendix C (online).

The final observation is that choosing a policy position less than $\lambda\tau$ away from r_1^* in the second period generates some gains in vote total compared to what R would have gotten with said position in period 1. Indeed, because of the compression of issue voters' ideal points, a small move to the left (right) of r_1^* doesn't lose votes on the outside (inside). However, because the initial density is singled-peaked, the density of the compressed second period distribution in the relevant interval will also be single-peaked. Thus, moving slightly to the right of r_1^* generates higher gains (compared to the vote total in period 1) than moving slightly to the left of r_1^* .

Put together, these observations imply that for any position D chooses in period 2 that is at most $\lambda\tau$ to the left or right of d_1^* , R 's best response will be slightly to the right of r_1^* . Thus, both parties have incentives to move slightly to the outside flank of their first period equilibrium positions. These incentives stop at $r_2^* = r_1^* + \lambda\tau$ and $d_2^* = -r_2^*$, for any additional move to the outside flank loses more votes on the inside to alienation than the additional votes gained on the outside. \square

Proof of Proposition 5. Suppose that $\lambda > \lambda^*$. Recall that in this case the equilibrium in the first period election consists of policy platforms $r_1^* \in (0, R)$ and $d_1^* = -r_1^*$ such that there is abstention only in the extremes; i.e., $\lambda > r_1^*$ (cf. Proposition A.4). As in the previous case, we implicitly assume that party R 's ideal point is sufficiently large and argue that its equilibrium position in the second period election is $r_2^* = r_1^*\tau + \lambda$.

We note that the observations made in the proof of Proposition 4 are still valid, subject to the obvious accommodations due to the fact that R 's marginal voters in the second election compressed differently: the left marginal voters situated at 0 compressed by $r_1^*\tau$; the right marginal voter situated at $r_1^* + \lambda$ compressed by $\lambda\tau > r_1^*\tau$.

Thus, it follows that in the second election both parties have incentives to move slightly to the outside flank of their first period equilibrium positions. These incentives stop at $r_2^* = r_1^*\tau + \lambda$

and $d_2^* = -r_2^*$. At these positions, both parties win with half probability. Any additional move to the outside flank loses more votes on the inside to alienation than the votes gained on the outside, again because the density of the second period election is single-peaked in the relevant intervals.

It remains to show that neither party has an incentive to jump back to the middle as long as $\lambda^* < \lambda \leq \bar{\lambda}$ and $\tau \geq \bar{\tau}$, for some $\bar{\lambda} > \lambda^*$ and $\bar{\tau} \in (0, 1)$. To do so, let $\bar{\lambda}$ and $\bar{\tau}$ be parameter values such that party R 's best response to $d = d_1^* \bar{\tau} - \bar{\lambda}$ in the first period election is $\bar{r} > 0$ satisfying $\bar{r} - \bar{\lambda} = d_1^* \bar{\tau}$. From [Proposition A.4](#) and [Lemma C.1](#), we know that $r_1^* < \bar{r} < r_1^* \bar{\tau} + \bar{\lambda}$. In this case, R has no incentive to jump to the middle. Indeed, any change from the second period equilibrium position $r_1^* \bar{\tau} + \bar{\lambda}$ would need to be a move to the left in order to gain vote share. However, because of the gap around 0 in the second period distribution of voters, such move would be need to be to a position at or left of \bar{r} . At this point, the vote share for R in the second election is the same as the vote share in the first selection, and thus party R does not move left of \bar{r} (which is by definition the first election best response to $d_1^* \bar{\tau} - \bar{\lambda}$). But R has the same vote share at \bar{r} and at $r_1^* \bar{\tau} + \bar{\lambda}$, the equilibrium position in the second period election. Thus R has no incentive to deviate.

The above argument remains valid as long as $\bar{\tau} \leq \tau < 1$ and $\lambda^* < \lambda \leq \bar{\lambda}$ because the first election best response function of R is a strictly convex function (see the proof of [Lemma C.1](#) in [Appendix C](#)). When $\lambda > \bar{\lambda}$ and $0 < \tau < \bar{\tau}$, party R has an incentive to jump from the position $r_1^* \tau + \lambda$ to what is the best response to position $d_1^* \tau - \lambda$ in the the first period election. In such case, the symmetric profile given by $r_2^* = r_1^* \tau + \lambda$ and $d_2^* = -r_2^*$ is no longer an equilibrium. In this case no symmetric equilibrium exists. If $-d_2 = r_2 > r_2^*$, then either party has incentive to get closer to the middle. If $-d_2 = r_2 < r_2^*$, then either party has incentive to move slightly to the outer flank. \square

Proof of Proposition 6. This follows readily from [Proposition 4](#) and [Proposition 5](#). \square

C Online Appendix

C.1 Hyperbolic Trigonometric Functions

The *hyperbolic cosine* function is defined on the real line by $\cosh(x) = \frac{1}{2}(e^x + e^{-x})$ and has range $[1, +\infty)$. One can show that it is an even function; i.e., $\cosh(-x) = \cosh(x)$ for all $x \in \mathbb{R}$, with $\cosh(0) = 1$.

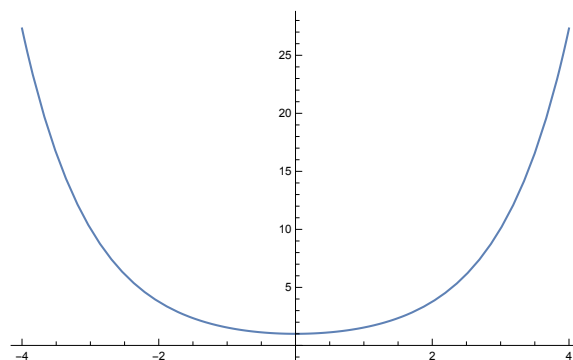


Figure 14: $\cosh(x)$.

The *hyperbolic sine* function is defined on the real line by $\sinh(x) = \frac{1}{2}(e^x - e^{-x})$ and has range $(-\infty, +\infty)$. One can show that it is an odd function; i.e., $\sinh(-x) = -\sinh(x)$ for all $x \in \mathbb{R}$, with $\sinh(0) = 0$.

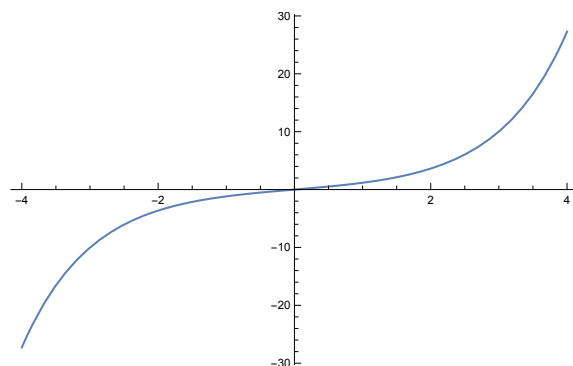


Figure 15: $\sinh(x)$.

The *hyperbolic tangent* function is defined by $\tanh(x) = \sinh(x) / \cosh(x)$. One can show that its range is the interval $(-1, 1)$, and further that this function is increasing.

Hyperbolic trigonometric functions satisfy some important properties, some of which are gathered in the next result.

Theorem C.1. *The following statements hold for the hyperbolic trigonometric functions:*

(a) $\frac{d}{dx} \cosh(x) = \sinh(x)$, $\frac{d}{dx} \sinh(x) = \cosh(x)$, and $\frac{d}{dx} \tanh(x) = 1 - \tanh^2(x)$.

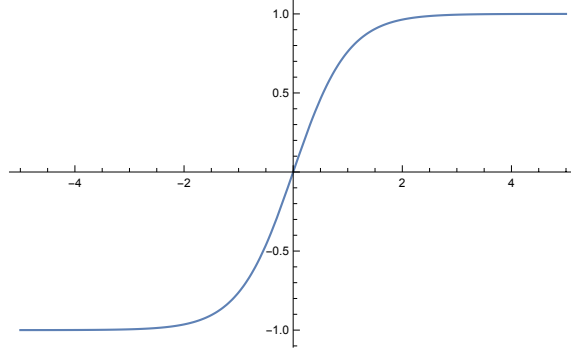


Figure 16: $\tanh(x)$.

- (b) $\cosh(x)$ is strictly convex on \mathbb{R} ;
- (c) $\sinh(x)$ is strictly convex on \mathbb{R}_- and strictly concave on \mathbb{R}_+ ;
- (d) $\tanh(x)$ is increasing on \mathbb{R} , strictly convex on \mathbb{R}_- and strictly concave on \mathbb{R}_+ ;
- (e) $\cosh^2(x) - \sinh^2(x) = 1$, for all $x \in \mathbb{R}$;
- (f) $\sinh(x) + \sinh(y) = 2 \sinh\left(\frac{x+y}{2}\right) \cosh\left(\frac{x-y}{2}\right)$;
- (g) $\cosh(x) + \cosh(y) = 2 \cosh\left(\frac{x+y}{2}\right) \cosh\left(\frac{x-y}{2}\right)$;
- (h) $\sinh(x+y) = \sinh(x) \cosh(y) + \cosh(x) \sinh(y)$;
- (i) $\cosh(x+y) = \cosh(x) \cosh(y) + \sinh(x) \sinh(y)$;
- (j) $\sinh(x-y) = \sinh(x) \cosh(y) - \cosh(x) \sinh(y)$;
- (k) $\cosh(x-y) = \cosh(x) \cosh(y) - \sinh(x) \sinh(y)$.

Proof. All the statements follow from the definitions of hyperbolic trigonometric functions. These are well-known results in the study of transcendental functions. For further reference, see [Roy and Olver \(2010\)](#). \square

C.2 Omitted Proofs in Appendix A

Proof of Lemma A.1. (a) Let $r \in [d, d+2\lambda] \cap \mathcal{P}$. We write the numerator of the vote share function for regime A as

$$F(r+\lambda; \alpha) - F\left(\frac{r+d}{2}; \alpha\right) = \frac{1}{1 + e^{-(r+\lambda)/\alpha}} - \frac{1}{1 + e^{-(r+d)/2\alpha}} = \frac{e^{-(r+d)/2\alpha} - e^{-(r+\lambda)/\alpha}}{(1 + e^{-(r+\lambda)/\alpha})(1 + e^{-(r+d)/2\alpha})}.$$

Similarly, we write the denominator of $S_A^R(r, d; \lambda, \alpha)$ as

$$F(r+\lambda; \alpha) - F(d-\lambda; \alpha) = \frac{1}{1 + e^{-(r+\lambda)/\alpha}} - \frac{1}{1 + e^{-(d-\lambda)/\alpha}} = \frac{e^{-(d-\lambda)/\alpha} - e^{-(r+\lambda)/\alpha}}{(1 + e^{-(r+\lambda)/\alpha})(1 + e^{-(d-\lambda)/\alpha})}.$$

Thus, we obtain

$$\begin{aligned}
S_A^R(r, d; \lambda, \alpha) &= \frac{e^{-(r+d)/2\alpha} - e^{-(r+\lambda)/\alpha}}{e^{-(d-\lambda)/\alpha} - e^{-(r+\lambda)/\alpha}} \times \frac{1 + e^{-(d-\lambda)/\alpha}}{1 + e^{-(r+d)/2\alpha}} \\
&= \left(1 + \frac{e^{(r+\lambda)/2\alpha}}{e^{(d-\lambda)/2\alpha}} \right)^{-1} \times \frac{1 + e^{-(d-\lambda)/\alpha}}{1 + e^{-(r+d)/2\alpha}} \\
&= \frac{1 + e^{-(d-\lambda)/\alpha}}{(1 + e^{-(r+d)/2\alpha})(1 + e^{(r-d+2\lambda)/2\alpha})} \\
&= \left(1 + \frac{e^{(r+\lambda)/2\alpha} + e^{-(r+\lambda)/2\alpha}}{e^{(d-\lambda)/2\alpha} + e^{-(d-\lambda)/2\alpha}} \right)^{-1} = \left(1 + \frac{\cosh\left(\frac{r+\lambda}{2\alpha}\right)}{\cosh\left(\frac{d-\lambda}{2\alpha}\right)} \right)^{-1}.
\end{aligned}$$

(b) Following the same line of arguments exhibited in part (a), one shows that

$$\begin{aligned}
S_B^R(r, d; \lambda, \alpha) &= \left(2 + \frac{\cosh\left(\frac{r}{\alpha}\right) - \cosh\left(\frac{d}{\alpha}\right)}{\cosh\left(\frac{\lambda}{\alpha}\right) + \cosh\left(\frac{d}{\alpha}\right)} \right)^{-1} \\
&= \left(1 + \frac{\cosh\left(\frac{r}{\alpha}\right) + \cosh\left(\frac{\lambda}{\alpha}\right)}{\cosh\left(\frac{d}{\alpha}\right) + \cosh\left(\frac{\lambda}{\alpha}\right)} \right)^{-1} = \left(1 + \frac{\cosh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right)}{\cosh\left(\frac{d+\lambda}{2\alpha}\right) \cosh\left(\frac{d-\lambda}{2\alpha}\right)} \right)^{-1}.
\end{aligned}$$

were the last equality follows from identity (g) in [Theorem C.1](#). \square

Proof of Lemma A.3. (a) For all $d \in \mathcal{P}$ and $\lambda, \alpha > 0$, one has $0 < [\cosh(\frac{d-\lambda}{2\alpha})]^{-1} \leq 1$. Thus, we replace this expression with a constant $0 < \kappa \leq 1$, so that

$$S_A^R(r, d; \lambda, \alpha) = \left(1 + \kappa \cosh\left(\frac{r+\lambda}{2\alpha}\right) \right)^{-1},$$

therefore

$$\log S_A^R(r, d; \lambda, \alpha) = -\log \left[1 + \kappa \cosh\left(\frac{r+\lambda}{2\alpha}\right) \right].$$

Twice differentiating the above expression obtains

$$\begin{aligned}
\frac{\partial^2 \log S_A^R(r, d; \lambda, \alpha)}{\partial r^2} &= -\frac{1}{2\alpha} \left[\frac{1}{2\alpha} \frac{\kappa \cosh\left(\frac{r+\lambda}{2\alpha}\right)}{1 + \kappa \cosh\left(\frac{r+\lambda}{2\alpha}\right)} - \frac{1}{2\alpha} \frac{\kappa^2 \sinh^2\left(\frac{r+\lambda}{2\alpha}\right)}{(1 + \kappa \cosh\left(\frac{r+\lambda}{2\alpha}\right))^2} \right] \\
&= -\frac{1}{4\alpha} \frac{\kappa \cosh\left(\frac{r+\lambda}{2\alpha}\right) + \kappa^2 [\cosh^2\left(\frac{r+\lambda}{2\alpha}\right) - \sinh^2\left(\frac{r+\lambda}{2\alpha}\right)]}{[1 + \kappa \cosh\left(\frac{r+\lambda}{2\alpha}\right)]^2}.
\end{aligned}$$

From identity (e) in [Theorem C.1](#) one sees that $\cosh^2(x) - \sinh^2(x) = 1$. Thus, the sign of the above expression is negative. It follows that $\log S_A^R(\cdot, d; \lambda, \alpha)$ is a strictly concave function, and thus $S_A^R(\cdot, d; \lambda, \alpha)$ is strictly log-concave, as desired.

(b) As before, for all $d \in \mathcal{P}$ and $\lambda, \alpha > 0$, one has $0 < [\cosh(\frac{d+\lambda}{2\alpha}) \cosh(\frac{d-\lambda}{2\alpha})]^{-1} \leq 1$. Replace this expression with a constant $0 < \kappa' \leq 1$, so that we have

$$\log S_B^R(r, d; \lambda, \alpha) = -\log \left[1 + \kappa' \cosh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right) \right].$$

Differentiating with respect to r gives

$$\frac{\partial \log S_B^R(r, d; \lambda, \alpha)}{\partial r} = -\frac{\kappa'}{2\alpha} \frac{\sinh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right) + \cosh\left(\frac{r+\lambda}{2\alpha}\right) \sinh\left(\frac{r-\lambda}{2\alpha}\right)}{1 + \kappa' \cosh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right)}.$$

Applying identity (h) in [Theorem C.1](#) on its numerator, we reduce the above expression to

$$\frac{\partial \log S_B^R(r, d; \lambda, \alpha)}{\partial r} = -\frac{\kappa'}{2\alpha} \frac{\sinh\left(\frac{r}{\alpha}\right)}{1 + \kappa' \cosh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right)}.$$

Differentiating with respect to r again obtains

$$\frac{\partial^2 \log S_B^R(r, d; \lambda, \alpha)}{\partial r^2} = -\frac{\kappa'}{2\alpha^2} \left[\frac{\cosh\left(\frac{r}{\alpha}\right)}{1 + \kappa' \cosh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right)} - \frac{\kappa'}{2} \frac{\sinh^2\left(\frac{r}{\alpha}\right)}{[1 + \kappa' \cosh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right)]^2} \right].$$

Let $\kappa'' = 1 + \kappa' \cosh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right)$. We now can write

$$\begin{aligned} \frac{\partial^2 \log S_B^R(r, d; \lambda, \alpha)}{\partial r^2} &= -\frac{\kappa'}{4\alpha^2} \left[\frac{2 \cosh\left(\frac{r}{\alpha}\right)}{\kappa''} - \frac{\kappa' \sinh^2\left(\frac{r}{\alpha}\right)}{(\kappa'')^2} \right] \\ &= -\frac{\kappa'}{4\alpha^2} \left[\frac{\cosh\left(\frac{r}{\alpha}\right) [2 + 2\kappa' \cosh\left(\frac{r+\lambda}{2\alpha}\right) \cosh\left(\frac{r-\lambda}{2\alpha}\right)] - \kappa' \sinh^2\left(\frac{r}{\alpha}\right)}{(\kappa'')^2} \right]. \end{aligned}$$

Applying identity (g) of [Theorem C.1](#) above obtains

$$\frac{\partial^2 \log S_B^R(r, d; \lambda, \alpha)}{\partial r^2} = -\frac{\kappa'}{4\alpha^2} \left[\frac{2 \cosh\left(\frac{r}{\alpha}\right) + \kappa' \cosh\left(\frac{r}{\alpha}\right) \cosh\left(\frac{\lambda}{\alpha}\right) + \kappa' (\cosh^2\left(\frac{r}{\alpha}\right) - \sinh^2\left(\frac{r}{\alpha}\right))}{(\kappa'')^2} \right].$$

Finally, using property (e) of [Theorem C.1](#) allows us to reduce the above expression to

$$\frac{\partial^2 \log S_B^R(r, d; \lambda, \alpha)}{\partial r^2} = -\frac{\kappa'}{4\alpha^2} \left[\frac{2 \cosh\left(\frac{r}{\alpha}\right) + \kappa' \cosh\left(\frac{r}{\alpha}\right) \cosh\left(\frac{\lambda}{\alpha}\right) + \kappa'}{(\kappa'')^2} \right].$$

Since $\cosh(x) \geq 1$ and $\kappa' > 0$, the sign of the above expression is negative. It follows that $\log S_B^R(\cdot, d; \lambda, \alpha)$ is strictly concave, thus $S_B^R(\cdot, d; \lambda, \alpha)$ is strictly log-concave. \square

Proof of Proposition A.2 (cont.) We show that R 's best response function in regime A is a contraction. Since voters are symmetrically distributed around zero and parties have opposite, but symmetric preferences, this means that D 's best response function is also a contraction. Existence and uniqueness of the result follows from the contraction mapping theorem.

To somewhat simplify the notation, we assume that $\alpha = 1$ and $\beta = 0$, and omit these terms on the expressions below. The analysis doesn't change with arbitrary $\alpha > 0$ and $\beta \geq 0$ as long as $\beta < 2\alpha$. The slope of R 's best response function in regime A given in Equation 11 is clearly zero in a corner solution. We first show that the slope of the interior solution $r_A^\circ(d; \lambda)$ satisfies

$$\left| \frac{\partial r_A^\circ(d; \lambda)}{\partial d} \right| < 1.$$

The FOC for an interior solution in regime A are

$$\mathcal{G}(r, d; \lambda) \equiv 1 - \frac{r - d}{2} \frac{\sinh\left(\frac{r+\lambda}{2}\right)}{\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right)} = 0.$$

By the implicit function theorem, we have

$$\frac{\partial r_A^\circ(d; \lambda)}{\partial d} = - \frac{\mathcal{G}_d(r, d; \lambda)}{\mathcal{G}_r(r, d; \lambda)},$$

where we use the subscripts in $\mathcal{G}(r, d; \lambda)$ to denote partial derivatives.

Partially differentiating \mathcal{G} with respect to d gives

$$\begin{aligned} \mathcal{G}_d(r, d; \lambda) &= \frac{1}{2} \frac{\sinh\left(\frac{r+\lambda}{2}\right)}{\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right)} + \frac{r - d}{4} \frac{\sinh\left(\frac{r+\lambda}{2}\right) \sinh\left(\frac{d-\lambda}{2}\right)}{[\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right)]^2} \\ &= \frac{\sinh\left(\frac{r+\lambda}{2}\right) \left\{ 2 \cosh\left(\frac{r+\lambda}{2}\right) + 2 \cosh\left(\frac{d-\lambda}{2}\right) + (r - d) \sinh\left(\frac{d-\lambda}{2}\right) \right\}}{4 [\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right)]^2}. \end{aligned}$$

Partially differentiating \mathcal{G} with respect to r gives

$$\begin{aligned} \mathcal{G}_r(r, d; \lambda) &= -\frac{1}{2} \frac{\sinh\left(\frac{r+\lambda}{2}\right)}{\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right)} \\ &\quad - \frac{r - d}{4} \left\{ \frac{\cosh\left(\frac{r+\lambda}{2}\right)}{\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right)} - \frac{\sinh^2\left(\frac{r+\lambda}{2}\right)}{[\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right)]^2} \right\} \\ &= -\frac{1}{2} \frac{\sinh\left(\frac{r+\lambda}{2}\right)}{\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right)} - \frac{r - d}{4 [\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right)]^2} \\ &\quad \times \left\{ \cosh^2\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{r+\lambda}{2}\right) \cosh\left(\frac{d-\lambda}{2}\right) - \sinh^2\left(\frac{r+\lambda}{2}\right) \right\} \\ &= \frac{2 \sinh\left(\frac{r+\lambda}{2}\right) [\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right)] + (r - d) \left(1 + \cosh\left(\frac{r+\lambda}{2}\right) \cosh\left(\frac{d-\lambda}{2}\right) \right)}{4 [\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right)]^2}. \end{aligned}$$

Combining these results and simplifying obtains

$$\begin{aligned}\frac{\partial r_A^\circ(d; \lambda)}{\partial d} &= \frac{2 \sinh\left(\frac{r+\lambda}{2}\right) \left[\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right) \right] + (r-d) \sinh\left(\frac{r+\lambda}{2}\right) \sinh\left(\frac{d-\lambda}{2}\right)}{2 \sinh\left(\frac{r+\lambda}{2}\right) \left[\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right) \right] + (r-d) \left[1 + \cosh\left(\frac{r+\lambda}{2}\right) \cosh\left(\frac{d-\lambda}{2}\right) \right]} \\ &= \frac{\varphi_A^{\text{num}}(r, d; \lambda)}{\varphi_A^{\text{den}}(r, d; \lambda)}.\end{aligned}\tag{16}$$

We have $d < r$ and also $r + \lambda > 0$. The first inequality follows from [Lemma A.4](#). The second from [Equation 10](#). Indeed, with $\alpha = 1$ and $\beta = 0$, if $r + \lambda < 0$ the FOC cannot be satisfied. Since $\cosh(x) \geq 1$ for all $x \in \mathbb{R}$ and $\sinh(x) \geq 0$ for $x \geq 0$, we obtain that $\varphi_A^{\text{den}}(r, d; \lambda)$ is always positive.

Subtract the numerator from the denominator of above expression to obtain

$$\begin{aligned}\varphi_A^{\text{den}}(r, d; \lambda) - \varphi_A^{\text{num}}(r, d; \lambda) &= (r-d) \left\{ 1 + \cosh\left(\frac{r+\lambda}{2}\right) \cosh\left(\frac{d-\lambda}{2}\right) - \sinh\left(\frac{r+\lambda}{2}\right) \sinh\left(\frac{d-\lambda}{2}\right) \right\} \\ &= (r-d) \left\{ 1 + \cosh\left(\frac{r-d+2\lambda}{2}\right) \right\} > 0,\end{aligned}$$

where the second equality follows from identity (k) in [Theorem C.1](#). Thus, the denominator of $\partial r_A^\circ(d; \lambda)/\partial d$ is strictly larger than its numerator. Since the denominator is strictly positive, it follows that

$$\partial r_A^\circ(d; \lambda)/\partial d < 1.\tag{17}$$

Now add the denominator and numerator of $\partial r_A^\circ(d; \lambda)/\partial d$ to obtain

$$\begin{aligned}\varphi_A^{\text{den}}(r, d; \lambda) + \varphi_A^{\text{num}}(r, d; \lambda) &= 4 \sinh\left(\frac{r+\lambda}{2}\right) \left[\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right) \right] \\ &\quad + (r-d) \left\{ 1 + \cosh\left(\frac{r+\lambda}{2}\right) \cosh\left(\frac{d-\lambda}{2}\right) + \sinh\left(\frac{r+\lambda}{2}\right) \sinh\left(\frac{d-\lambda}{2}\right) \right\} \\ &= 4 \sinh\left(\frac{r+\lambda}{2}\right) \left[\cosh\left(\frac{r+\lambda}{2}\right) + \cosh\left(\frac{d-\lambda}{2}\right) \right] \\ &\quad + (r-d) \left\{ 1 + \cosh\left(\frac{r+d}{2}\right) \right\} > 0,\end{aligned}$$

where the second equality follows from identity (i) in [Theorem C.1](#). Since the denominator is strictly positive, we can rearrange this expression to obtain $\varphi_A^{\text{num}}(r, d; \lambda)/\varphi_A^{\text{den}}(r, d; \lambda) > -1$. It follows that

$$-1 < \partial r_A^\circ(d; \lambda)/\partial d.\tag{18}$$

Combining [Equation 17](#) and [Equation 18](#) obtains

$$-1 < \partial r_A^\circ(d; \lambda)/\partial d < 1.$$

Thus, R 's interior best response function, defined for $d \in [D, R]$ has a slope whose absolute value is strictly less than 1. To show that it is a contraction, note that $\partial r_A^\circ(d; \lambda)/\partial d$ in [Equation 16](#) is expressed in terms of $\sinh(x)$ and $\cosh(x)$, and thus is continuous. This implies that $|\partial r_A^\circ(d; \lambda)/\partial d|$

is continuous on $[D, R]$, and hence has a maximizer. This provides a uniform bound on the absolute value of the slope of R 's best response function. \square

Proof of Proposition A.3 (cont.) We show that the slope of the interior solution $r_B^\circ(d; \lambda, \alpha, \beta)$ satisfies

$$\left| \frac{\partial r_B^\circ(d; \lambda, \alpha, \beta)}{\partial d} \right| < 1.$$

The rest of the argument is similar to the one used in the proof of Proposition A.2. To somewhat simplify the notation, we assume that $\alpha = 1$ and $\beta = 0$, and omit these terms on the expressions below. As in the previous case, the analysis doesn't change with arbitrary $\alpha > 0$ and $\beta \geq 0$.

The FOC in regime B are

$$\mathcal{H}(r, d; \lambda) \equiv 1 - \frac{r-d}{2} \frac{\sinh(r)}{\cosh\left(\frac{r+\lambda}{2}\right) \cosh\left(\frac{r-\lambda}{2}\right) + \cosh\left(\frac{d+\lambda}{2}\right) \cosh\left(\frac{d-\lambda}{2}\right)} = 0.$$

By the implicit function theorem, we have that

$$\frac{\partial r_B^\circ(d; \lambda)}{\partial d} = - \frac{\mathcal{H}_d(r, d; \lambda)}{\mathcal{H}_r(r, d; \lambda)}.$$

Partially differentiating \mathcal{H} with respect to d gives

$$\begin{aligned} \mathcal{H}_d(r, d; \lambda) &= \frac{\sinh r \left[2 \cosh \frac{r-\lambda}{2} + (r-d) \cosh \frac{d+\lambda}{2} \sinh \frac{d-\lambda}{2} + \cosh \frac{d-\lambda}{2} \left(2 \cosh \frac{d+\lambda}{2} + (r-d) \sinh \frac{d+\lambda}{2} \right) \right]}{4 \left[\cosh \frac{r+\lambda}{2} \cosh \frac{r-\lambda}{2} + \cosh \frac{d+\lambda}{2} \cosh \frac{d-\lambda}{2} \right]^2}. \end{aligned}$$

Partially differentiation \mathcal{H} with respect to r gives

$$\begin{aligned} \mathcal{H}_r(r, d; \lambda) &= (r-d) \cosh\left(\frac{r+\lambda}{2}\right) \sinh r \sinh\left(\frac{r-\lambda}{2}\right) - 2 \cosh\left(\frac{d-\lambda}{2}\right) \cosh\left(\frac{d+\lambda}{2}\right) [\sinh r + (r-d) \cosh r] \\ &\quad + \cosh\left(\frac{r-\lambda}{2}\right) \{ (r-d) \sinh r \sinh\left(\frac{r+\lambda}{2}\right) - 2 \cosh\left(\frac{r+\lambda}{2}\right) [\sinh r + (r-d) \cosh r] \} / \\ &\quad 4 \left[\cosh\left(\frac{r+\lambda}{2}\right) \cosh\left(\frac{r-\lambda}{2}\right) + \cosh\left(\frac{d+\lambda}{2}\right) \cosh\left(\frac{d-\lambda}{2}\right) \right]^2. \end{aligned}$$

\mathcal{H}_d and \mathcal{H}_r have the same denominator. Using identities (g) and (h) of Theorem C.1 allows us to simplify the numerator of $\mathcal{H}_d(r, d; \lambda)$ to obtain

$$\phi_B^{\text{num}}(r, d; \lambda) = \sinh(r) [\cosh(r) + \cosh(d) + 2 \cosh(\lambda) + (r-d) \sinh(d)].$$

Factorizing and using identities (g) and (h) again to simplify the negative of the numerator of $\mathcal{H}_r(r, d; \lambda)$ obtains

$$\phi_B^{\text{den}}(r, d; \lambda) = [\sinh(r) + (r-d) \cosh(r)] [\cosh(d) + 2 \cosh(\lambda)] + (r-d) + \sinh(r) \cosh(r).$$

We have

$$\frac{\partial r_B^\circ(d; \lambda)}{\partial d} = - \frac{\mathcal{H}_d(r, d; \lambda)}{\mathcal{H}_r(r, d; \lambda)} = \frac{\varphi_B^{\text{num}}(r, d; \lambda)}{\varphi_B^{\text{den}}(r, d; \lambda)}. \quad (19)$$

Note that both numerator and denominator are non-negative for $r \geq 0$ and $d \leq 0$. Subtract the denominator from the numerator of the above expression to obtain

$$\begin{aligned} \varphi_B^{\text{num}}(r, d; \lambda) - \varphi_B^{\text{den}}(r, d; \lambda) &= (r - d) \times \\ &\quad [\sinh(r) \sinh(d) - \cosh(r) \cosh(d) - 2 \cosh(r) \cosh(\lambda) - 1] \\ &< 0, \end{aligned}$$

where the last inequality follows from $\sinh(r) \geq 0$ and $\sinh(d) \leq 0$. Thus, it follows that

$$\partial r_B^\circ(d; \lambda) / \partial d < 1. \quad (20)$$

Now add the denominator to the numerator to obtain

$$\begin{aligned} \varphi_B^{\text{num}}(r, d; \lambda) + \varphi_B^{\text{den}}(r, d; \lambda) &= 2 \sinh(r) [\cosh(r) + 2 \cosh(\lambda) + \cosh(d)] \\ &\quad + (r - d) [\cosh(r) \cosh(d) + \sinh(r) \sinh(d) + 1] \\ &= 2 \sinh(r) [\cosh r + 2 \cosh \lambda + \cosh d] \\ &\quad + (r - d) [\cosh(r + d) + 1] \\ &> 0, \end{aligned}$$

where the last equality follows from identity (i) in [Theorem C.1](#). Since the denominator is strictly positive, we can rearrange this expression to obtain $\varphi_B^{\text{num}}(r, d; \lambda) / \varphi_B^{\text{den}}(r, d; \lambda) > -1$. It follows that

$$-1 < \partial r_B^\circ(d; \lambda) / \partial d. \quad (21)$$

Combining [Equation 20](#) and [Equation 21](#) obtains

$$-1 < \partial r_B^\circ(d; \lambda) / \partial d < 1.$$

The rest follows the arguments at the end of the proof of [Proposition A.2](#). □

C.3 Additional Derivations for the First Period Election

Lemma C.1. *Let $\lambda, \alpha > 0$ and $0 \leq \beta < 2\alpha$ be given.*

- (a) *Party R's interior best response function $r_A^\circ(d; \lambda, \alpha, \beta)$ in regime A attains a unique minimum at $d = -r_A^\circ(\lambda, \alpha, \beta)$, where $r_A^\circ(\lambda, \alpha, \beta)$ is R's interior symmetric equilibrium strategy implicitly defined by [Equation 12](#).*
- (b) *Party R's (interior) best response function $r_B^\circ(d; \lambda, \alpha, \beta)$ in regime B attains a unique minimum at $d = -r_B^\circ(\lambda, \alpha, \beta)$, where $r_B^\circ(\lambda, \alpha, \beta)$ is R's interior symmetric equilibrium strategy implicitly defined by [Equation 15](#).*

Proof. We let $\alpha = 1$ and $\beta = 0$ for notational convenience; our results do not depend on this choice of parameter values. We provide a proof for part (b). The proof of part (a) follows the same arguments.

(b) We first show that $r_B^\circ(d; \lambda)$ is minimized at $d = -r_B^\circ(\lambda)$, where using [Equation 15](#) this last is characterized by

$$1 = r_B^\circ(\lambda) \frac{\sinh r_B^\circ(\lambda)}{\cosh r_B^\circ(\lambda) + \cosh \lambda}.$$

[Equation 19](#) provides an expression for the partial derivative of $r_B^\circ(d; \lambda)$ with respect to d . Note that the denominator of this expression, $\varphi_B^{\text{den}}(r, d; \lambda)$, is never zero. Evaluating its numerator $\varphi_B^{\text{num}}(r, d; \lambda)$ at $r = -d = r_B^\circ(\lambda)$ yields

$$\begin{aligned} \varphi_B^{\text{num}}(r, d; \lambda) \Big|_{r=-d=r_B^\circ(\lambda)} &= 2 \sinh(r_B^\circ(\lambda)) [\cosh(r_B^\circ(\lambda)) + \cosh(\lambda) - r_B^\circ(\lambda) \sinh(r_B^\circ(\lambda))] \\ &= 0. \end{aligned}$$

Thus, it follows that

$$\frac{\partial r_B^\circ(d; \lambda)}{\partial d} \Big|_{r=-d=r_B^\circ(\lambda)} = 0.$$

It remains to show that $d = -r_B^\circ(\lambda)$ is indeed a minimizer. For that it suffices to argue that the function $r_B^\circ(d; \lambda)$ implicitly defined by [Equation 13](#) is convex.

We use the implicit function $\mathcal{H}(r, d; \lambda) = 0$, which is the FOC for an interior maximizer in regime B . Since both r and d are one dimensional, we have that $r_B^\circ(d; \lambda)$ is strictly convex in d if and only if

$$\mathcal{H}_d \mathcal{H}_{rd} - \mathcal{H}_r \mathcal{H}_{dd} > 0.$$

Taking the corresponding partial derivatives to \mathcal{H}_r and \mathcal{H}_d and simplifying yields

$$\begin{aligned} &\mathcal{H}_d \mathcal{H}_{rd} - \mathcal{H}_r \mathcal{H}_{dd} \\ &= \frac{1}{2(2 \cosh \lambda + \cosh r + \cosh d)^4} \times \sinh(r) \left\{ (3 + 2(r-d)^2) \cosh(d) + \cosh(d-2r) \right. \\ &\quad + \cosh(2d-r) + 4 \cosh^2(\lambda) \cosh(r) + (5 + (r-d)^2 + 2 \cosh(2\lambda)) \cosh r \\ &\quad + 4(r-d) \cosh^2\left(\frac{d-r}{2}\right) \sinh(r) + 2 \cosh(\lambda) \left(3 + 2[2 + (r-d)^2] \cosh(d) \cosh(r) \right. \\ &\quad \left. \left. + \cosh(2r) - 2 \underline{[(d-r) \cosh(d) + \sinh(d)] \sinh(r)} \right) \right\}. \end{aligned}$$

The denominator of the above expression is strictly positive. Since $r \geq 0$ by [Lemma A.2](#) and $\cosh(x) \geq 1$, one sees that the non-underlined component of the numerator is strictly positive.

Manipulating the underlined expression gives

$$\begin{aligned} -2[(d-r) \cosh(d) + \sinh(d)] \sinh(r) &= 2[(r-d) \cosh(d) - \sinh(d)] \sinh(r) \\ &= 2(r-d) \cosh(d) \sinh(r) - \sinh(d) \sinh(r) \\ &= 2(r-d) \cosh(d) \sinh(r) + \sinh(-d) \sinh(r). \end{aligned}$$

From [Theorem C.1](#) one has that $2(r - d) \cosh(d) \sinh(r) \geq 0$. Noting that $-d \geq 0$, we also have that $\sinh(-d) \sinh(r) \geq 0$. Therefore, it is true that

$$-2[(d - r) \cosh(d) + \sinh(d)] \sinh(r) \geq 0.$$

It follows that the numerator of the expression we are interested in is also strictly positive. Thus, we conclude that

$$\mathcal{H}_d \mathcal{H}_{rd} - \mathcal{H}_r \mathcal{H}_{dd} > 0,$$

as desired. Hence we have that $r_B^\circ(d; \lambda)$ is a strictly convex function, with a global minimum at $d = -r_B^\circ(\lambda)$. This concludes the proof of the lemma. \square

Lemma C.2. *Let $r_A^\circ(\lambda, \alpha, \beta)$ and $r_B^\circ(\lambda, \alpha, \beta)$ be implicitly defined by [Equation 12](#) and [Equation 15](#), respectively. One has that*

$$\frac{\partial r_A^\circ(\lambda, \alpha, \beta)}{\partial \lambda} < 0 \quad \text{and} \quad \frac{\partial r_B^\circ(\lambda, \alpha, \beta)}{\partial \lambda} > 0.$$

Proof. Using the implicit function theorem, one obtains from [Equation 12](#)

$$\frac{\partial r_A^\circ(\lambda, \alpha, \beta)}{\partial \lambda} = - \frac{\left(\frac{r}{2\alpha} + \frac{\beta}{4\alpha}\right) \partial[\tanh(\frac{r+\lambda}{2\alpha})]}{\tanh(\frac{r+\lambda}{2\alpha}) + \left(\frac{r}{2\alpha} + \frac{\beta}{4\alpha}\right) \partial[\tanh(\frac{r+\lambda}{2\alpha})]}.$$

Here $\partial[\tanh(\frac{r+\lambda}{2\alpha})]$ refers to the derivative of the $\tanh(x)$ function. Since this last is strictly positive, $\lambda > 0$, and $r_A^\circ > 0$ in the interior symmetric equilibrium for regime A, one has $\partial r_A^\circ(\lambda, \alpha, \beta)/\partial \lambda < 0$.

To show the sign of the partial derivative of r_B° , taking partial derivative with respect to λ to the right-hand side of [Equation 15](#) obtains

$$C_1 = -\frac{1}{2\alpha} \left(\frac{r}{2\alpha} + \frac{\beta}{4\alpha}\right) \left\{ \tanh^2\left(\frac{r-\lambda}{2\alpha}\right) - \tanh^2\left(\frac{r+\lambda}{2\alpha}\right) \right\} < 0,$$

where the last inequality follows from the fact that $\lambda > 0$ and the hyperbolic tangent function is an increasing function. Taking partial derivative with respect to r to the right-hand side of [Equation 15](#) obtains

$$C_2 = \frac{1}{2\alpha} \left\{ \tanh\left(\frac{r+\lambda}{2\alpha}\right) + \tanh\left(\frac{r-\lambda}{2\alpha}\right) \right\} + \frac{1}{2\alpha} \left(\frac{r}{2\alpha} + \frac{\beta}{4\alpha}\right) \left\{ 2 - \tanh^2\left(\frac{r+\lambda}{2\alpha}\right) - \tanh^2\left(\frac{r-\lambda}{2\alpha}\right) \right\}.$$

Since $\alpha, \lambda > 0$ and $r > \lambda$ in the interior symmetric equilibrium for regime B, and also $|\tanh(x)| \leq 1$ for all $x \in \mathbb{R}$, we have $C_2 > 0$. Thus, by the implicit function theorem, $\partial r_B^\circ(\lambda, \alpha, \beta)/\partial \lambda = -C_1/C_2 > 0$, as desired. \square