FEDERAL RESERVE BANK
PHILADELPHIA

Working Papers
RESEARCH DEPARTMENT

The Changing Polarization of Party Ideologies

The Role of Sorting

Satyajit Chatterjee
Federal Reserve Bank of Philadelphia Research Department

Burcu Eyigungor
Federal Reserve Bank of Philadelphia Research Department

WP 23-07

PUBLISHED
May 2023

ISSN: 1962-5361

Disclaimer: This Philadelphia Fed working paper represents preliminary research that is being circulated for discussion purposes. The views expressed in these papers are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Any errors or omissions are the responsibility of the authors. Philadelphia Fed working papers are free to download at: https://philadelphiafed.org/research-and-data/publications/working-papers.

DOI: https://doi.org/10.21799/frbp.wp.2023.07

The Changing Polarization of Party Ideologies: The Role of Sorting

Satyajit Chatterjee and Burcu Eyigungor

Federal Reserve Bank of Philadelphia

February 2023

Abstract

U.S. congressional roll-call voting records show that as polarization of the two parties along the economic dimension changes, polarization along the social/cultural dimension tends to change in the opposite direction. A model of party competition within a two-dimensional ideology space is developed in which party platforms are determined by voters who compose the party. It is shown that if distribution of voter preferences is radially symmetric, polarization of party ideologies along the two dimensions are inversely related, as observed. The model gives a remarkably good quantitative account of the historically observed movements in polarization along the two dimensions.

Keywords: polarization, primaries, partisan sorting, political economy
JEL Codes: D72 P16

---

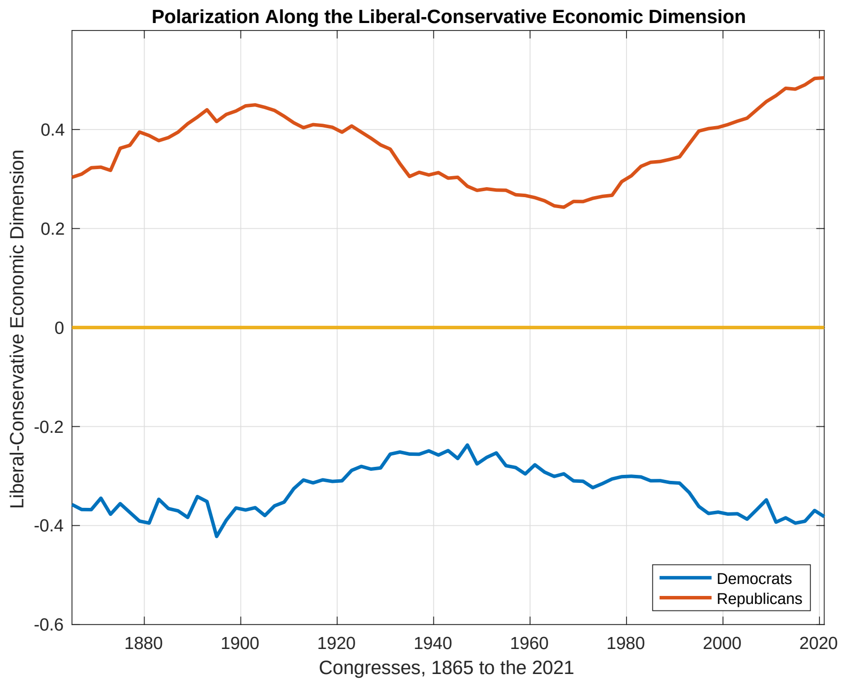
satyajit.chatterjee@phil.frb.org & burcu.eyigungor@phil.frb.org

Acknowledgements and Disclaimer: The authors thank conference and seminar participants at the 2022 SED Annual Meeting, 2022 Santiago Macro Workshop, UBC and ISI Delhi for thoughtful comments. The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Any errors or omissions are the responsibility of the authors. Philadelphia Fed working papers are free to download at https://philadelphiafed.org/research-and-data/publications/working-papers.

1 Introduction

Much of the analysis of the Poole and Rosenthal (1997) roll-call voting data has focused on the polarization of political parties along the liberal-conservative economic dimension. Figure 1 shows

Figure 1:
Polarization Along the Liberal-Conservative Dimension

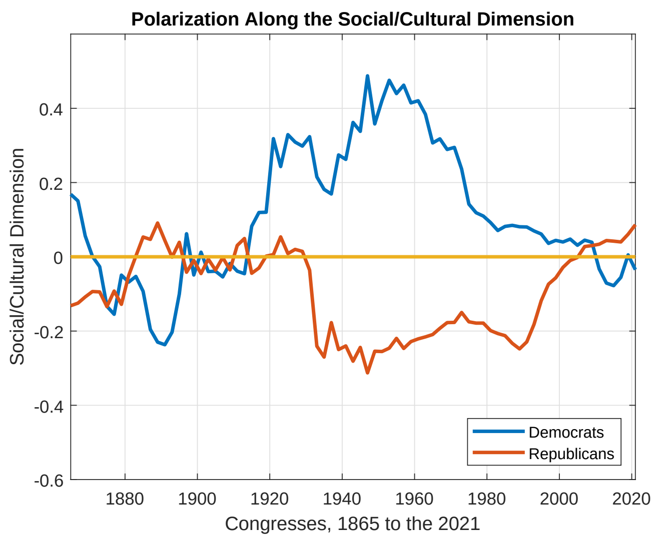


Source: Lewis, Poole, Rosenthal, Boche, Rudkin, and Sonnet (2022)

the evolution of polarization along this dimension since 1865. It plots the mean ideological position of Democratic and Republican representatives in Congress, from the 39th Congress to the (current) 117th Congress. Throughout this period, the Republican party has occupied the conservative end of the spectrum and the Democratic party the liberal end. That said, the ideological distance between the two parties has waxed and waned over time: The distance shrank during the first eight decades of the 20th century but has been rising since 1980. Currently, the distance is the widest it has been since 1865.

In contrast, the ideological distance along the second dimension has received less attention. We refer to the second dimension as the dimension of social/cultural issues, with positive numerical values indicating social conservatism and negative numerical values indicating social liberalism (Figure 2). At the end of the Civil War, the Democratic party was the socially conservative party, while the Republican party was socially liberal. However, this changed during the final decades of the 19th century as the Democratic party became socially liberal and the Republican party socially conservative. In the first couple of decades of the 20th century, there was hardly any difference between the two parties along this dimension. Between 1920 and 1960, the ideological difference widened, with Democrats becoming increasingly socially conservative. The trend reversed starting in the mid-1960s and the ideological distance between the two parties is currently low. That said, the Republican party is now the socially conservative party, while the Democratic party is the socially liberal one.

Figure 2:
Polarization Along the Social/Cultural Dimension

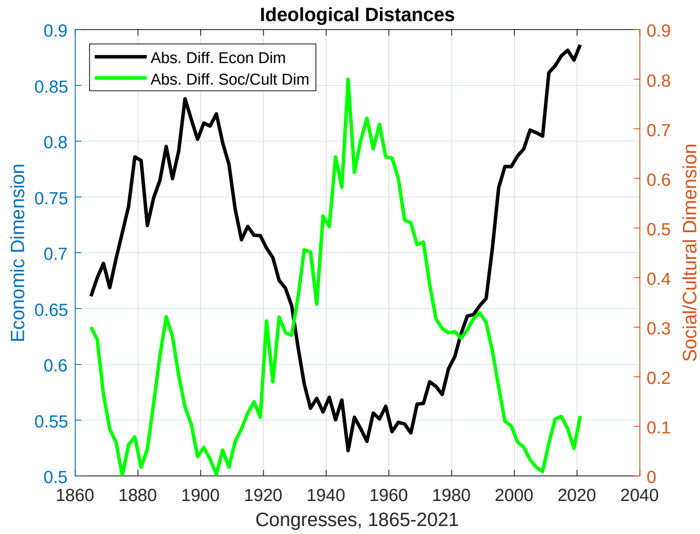


Source: Lewis, Poole, Rosenthal, Boche, Rudkin, and Sonnet (2022)

For the purposes of this paper, there are two key takeaways of these plots. The first is the manner in which these ideological distances move relative to each other: When ideological distance increases along one dimension, it tends to *decrease* along the other dimension. This pattern of inverse association is easily seen in time plots of the absolute difference in the ideological positions of the two parties for each dimension, as shown in Figure 3: Periods when ideological distance along the social/cultural dimension rose tend to be periods when the ideological distance along the economic dimension fell.[1] Indeed, the correlation between the two series is -0.86.

[1]The optimal classification (OC) technique (Poole (2005)) that locates a congress-person on the two-dimensional ideology space implies only that the person's ideological position must lie in the square $[-1, 1] \times [-1, 1]$. The inverse association shown in Figure 3 is not implied by the mechanics of OC but is, instead, a feature of the roll-call voting patterns.

Figure 3:
Ideological Distance Between Parties Along
Economic and Social Dimensions



Source: Lewis, Poole, Rosenthal, Boche, Rudkin, and Sonnet (2022) and authors' calculation

The second takeaway is that the ideological orientation of a party — liberal or conservative — along a dimension is not necessarily fixed. While the orientation of each party on economic issues has not changed since the Civil War, orientation on social issues has changed. The Democratic party currently espouses a liberal social ideology but for the better part of the post-Civil War period its representatives have, on average, voted conservatively on social issues. Similarly, the reverse holds for the Republican party.

The main contribution of this paper is to show that the negative association between the two distance measures shown in Figure 3 might not be a historical accident. In fact, it is implied by a spatial model of voting in which party ideologies are formed in primary elections that precede the national election.

In our model, the ideological stance of a party is determined by the composition of voters in a party's primary election. Voters are heterogeneous regarding their preferences (equivalently, ideologies) and sort into the two parties depending on what they believe each party stands for. The primary elections then turn the preferences of primary attendees into party ideology and platform for the national election. Importantly, the desire of each party to win the national election results in party platforms that trade off the preferences of party adherents against the mean preferences of the polity at large. It is this balancing act that accounts for Figure 3: When polarization along the social dimension increases, polarization along the economic dimension shrinks so as to keep the overall extremeness of a party's ideology — measured as the Euclidean distance of a party's ideology from the mean ideology of the polity — constant.

In our model, the changes in mean ideology of the two parties shown in Figures 1 and 2 is attributed entirely to re-sorting, i.e., to changes in the composition of voters making up the two parties. There is no change in the distribution of voter ideologies or in any other fundamental. What, then, triggers a change in the composition of voters making up a party? It is caused by a self-fulfilling change in voters' (common) beliefs about party ideologies. Specifically, if the distribution of voter ideologies is radially symmetric (a common assumption in spatial voting models which we also make), our model implies the existence of a circle in the two-dimensional ideology space such that any two diametrically opposite points on that circle constitute *equilibrium* beliefs. Meaning, if voters associate the two parties with those (diametrically) opposing ideologies, the resulting partisan sorting and competition for national power will generate party ideologies that reproduce those beliefs.

Although the stability of voter preferences over such a long span of time is surely not exactly true, the sorting and re-sorting of a stable distribution of voters into the two parties is consistent with the malleability of party ideologies evident in the figures. More importantly, we show that our model of belief-based partisan sorting gives a surprisingly good *quantitative* account of the evolution of ideological distances seen in Figures 1 and 2.

The paper is organized as follows. In Section 2, we present a brief discussion of the extant literature in economics and politics that relates to our paper. Section 3 lays out the model environment. Section 4 solves for equilibria that display polarization along one dimension (motivated by Figure 1) but may or may not display polarization along the other dimension (motivated by Figure 2). This section explains why there is an inverse association between the two ideological distances shown in Figure 3. Section 5 takes this insight further and shows that the variation in ideological distances shown in Figure 3 can be well-accounted for, in a quantitative sense, simply by the *direction* of party ideologies implicit in Figures 1 and 2. Section 6 concludes. The Appendices contain proofs of some propositions and other supplementary materials.

2 Connections to the Literature

Our paper contributes to the interrelated literatures on party competition and polarization. In Downsian models of party competition (Downs (1957)), both parties declare the same platform (convergence of policies/platforms). As pointed out by many, this prediction is at odds with the facts as parties typically offer different platforms and it is quite inconsistent with Figures 1 and 2 which show that the two US parties have consistently offered distinct ideologies.

Given this dissonance, an extensive theoretical literature has developed to explain policy divergence.[2] One strand of this literature, beginning with Wittman (1973), sought to locate policy divergence in exogenous differences in party preferences. More recently, Roemer (2001) proposed an extension of Wittman's model in which differences in party preferences are derived from the different preferences of voters who compose these parties. Our modeling approach is in this spirit, but our theory of how preferences of party members is aggregated to produce the party platform is different. We assume this preference aggregation occurs via primary elections while in Roemer (2001) it comes about through bargaining between party factions.

The rise in party polarization since the early 1980s was documented by Poole and Rosenthal (1997). Some papers that explain policy divergence also suggest reasons why polarization has increased in recent decades. When the focus is on a single policy dimension, some set of external factors must change to account for increasing polarization. Along these lines, the literature has pointed to several possible causes.

McCarty, Poole, and Rosenthal (2006) focus on the fact that the two parties are currently much more stratified by income than in the past: Higher income individuals identify with the Republican party (and lower income individuals with the Democratic party) more so now than 50 years ago. They view this development as the likely cause of increasing polarization of the two parties.[3]

A second force that has attracted attention are campaign contributions. Herrera, Levine, and Martinelli (2008) present a model in which increasing uncertainty about election outcomes increases the size of campaign contributions as well as polarization of party platforms. Drautzburg, Livshits, and Wright (2021) show how increases in legal contribution limits to political campaigns can lead

[2] See De Donder and Gallego (2017) for a recent survey of the literature on the positioning of political parties in unidimensional and multidimensional policy spaces.

[3] See Barber and McCarty (2013) for a recent survey of the more empirical research on the causes and consequences of polarization.

to greater polarization in candidate positions, provided campaign contributions generate private benefits for contestants. Konishi and Pan (2020) study a model in which policy choices have two dimensions: social and economic. In their model, third parties lobby *both* contestants to ensure the economic dimension is not politically salient. This leaves contestants to differentiate themselves by espousing different social policies, and divergence along this dimension can increase if lobbying on the economic dimension becomes more intense.

A third force that has attracted attention is the role of imperfect understanding of the true state of affairs. Dixit and Weibull (2007) present a model in which policy failures lead some individuals to want to reverse course and others to advocate an even stronger form of the failed policy. Along similar lines, although not linked directly to political polarization, Azzimonti and Fernandes (2021) present a model in which the advent of social media as a news source causes polarization of beliefs.

In contrast to these explanations, our paper connects changes in polarization along the economic dimension to changes in polarization along the social/cultural dimension (and vice versa). Similar to us, Krasa and Polborn (2014) have examined how polarization along the economic dimension is affected by polarization along the social/cultural dimension. They present a model in which contestants have opposing views on cultural issues but choose positions along the economic dimension. Under certain conditions, increased divergence between the candidates along the social dimension leads to increased polarization along the economic dimension. Our approach differs from theirs in that a party does not inherit the views of the candidate representing it, rather, candidates commit to the platform/ideology that primary attendees collectively want them to espouse. And, consistent with the inverse association between the ideological distances along the social and economic dimensions documented earlier, our model predicts that economic polarization rises when social polarization *falls*.

Political polarization has also attracted the attention of macroeconomists. This literature assumes two types of agents, each represented by a political party, where the probability of a party becoming the ruling party is either given or determined in a general election.[4] This framework has been used to study the impact of polarized preferences on fiscal policy, in Alesina and Tabellini (1990); on aggregate investment, in Azzimonti (2011); on sovereign borrowing and default, in Cuadra and Sapriza (2008); on allocative inefficiencies, in Acemoglu, Golosov, and Tsyvinski (2011);

[4]For a closely related literature that models fiscal policy as the outcome of legislative bargaining game, see Battaglini and Coate (2008)

on policy extremism and reelection probabilities, in Chatterjee and Eyigungor (2020). All these studies are explicitly dynamic and most are quantitative. Our paper shares the quantitative focus of this literature but abstracts from explicit dynamics in order to delve deeper into the formation of party preferences.

3 The Environment

3.1 Voters

There is a set of *voters* who care about two issues. In the space of issues, a voter's *type* is defined by $(x, y) \in \mathbb{R}^2$. If policies chosen by the government is (w, z), the utility of a type (x, y) voter is:

$$U(x, y; w, z) = -(w - x)^2 - (z - y)^2. \tag{1}$$

Thus (x, y) denotes the voter's ideal policies. The density of the distribution of voter types is $f(x, y) \geq 0$. We assume that this distribution is symmetric around $(0, 0)$, i.e.,

$$f(x, y) = f(-x, -y) \quad \forall\, (x, y) \in \mathbb{R}^2. \tag{2}$$

3.2 Political Parties

A *political party* is an entity characterized by its *policy platform*, i.e, by a pair $(w, z) \in \mathbb{R}^2$ of policies that the entity promises to implement if it gains the power to govern. We refer to the two parties as the D party and the R party and denote their respective platforms by (w_k, z_k), $k \in \{D, R\}$ and the combined party platforms by $\mathcal{P} = (w_D, z_D, w_R, z_R)$.

3.3 Electoral Institutions

The identity of the party that gets to govern is determined in a *national election* via majority vote. We assume that a voter of type (x, y) votes for the D party if and only if

$$U(x, y; w_D, z_D) + A \geq U(x, y; w_R, w_D). \tag{3}$$

Here $A \in \mathbb{R}$ is a net preference in favor of the D party for reasons unrelated to policies and is a random variable whose value is realized at the time of the election. It is drawn from a probability distribution that is symmetric around 0 and has a CDF

$$F(A) : \mathbb{R} \to [0,1] \text{ continuous in } A. \tag{4}$$

Given party platforms $\mathcal{P}$ and an A, let $\chi(x, y; A, \mathcal{P})$ be an indicator function that takes the value 1 if (3) is satisfied and 0 otherwise. Then $\mathbb{E}_{(x,y)}\chi(x, y; A, \mathcal{P})$ is the measure of the set of voters who vote for the D party. Let

$$\mathbb{A}(\mathcal{P}) \equiv \{A \in \mathbb{R} : \mathbb{E}_{(x,y)}\chi(x, y; A, \mathcal{P}) \geq 0.5\} \tag{5}$$

be the set of A values for which the measure of people who vote for the D party is at least one-half, given $\mathcal{P}$. Then

$$\pi(\mathcal{P}) = \mathbb{E}_A \mathbb{1}_{\{A \in \mathbb{A}(\mathcal{P})\}} \tag{6}$$

is the probability of D party winning the election, given $\mathcal{P}$.

The policy platforms of the two parties are determined simultaneously in two separate *primary elections* that precede the national election. In describing the outcome of these elections we rely on the well-known equivalence (Lindbeck and Weibull (1987)) between maximization of (utilitarian) social welfare and the equilibrium outcome of a probabilistic voting games in which the people voting in primary k choose between candidates who care only about winning the primary election (i.e., they are office motivated in the primary election).

To describe the determination of these platforms, let $\mathcal{P}^e = (w^e_D, z^e_D, w^e_R, z^e_R)$ denote the voters' (common) expectations of party platforms in the upcoming national election. Define $H_k(\mathcal{P}^e)$ as the set of voters who have a strict or a weak preference for platform (w^e_k, z^e_k). That is,

$$H_k(\mathcal{P}^e) = \{(x, y) \in \mathbb{R}^2 : U(x, y; w^e_k, z^e_k) \geq U(x, y; w^e_{\sim k}, z^e_{\sim k})\}, \tag{7}$$

where $\sim k = R$ (D) if $k = D$ (R). Then the D-party platform solves:

$$(w_D, z_D) = \underset{(w,z)}{\operatorname{argmax}} \left\{ \begin{array}{c} \pi(w, z, w_R, z_R)\, \mathbb{E}_{(x,y)} \left(\left[-(w - x)^2 - (z - y)^2 \right] \mathbb{1}_{\{(x,y) \in \mathbb{H}_D(\mathcal{P}^e)\}} \right) \\ + \left[1 - \pi(w, z, w_R, z_R)\right] \mathbb{E}_{(x,y)} \left(\left[-(w^e_R - x)^2 - (z^e_R - y)^2 \right] \mathbb{1}_{\{(x,y) \in \mathbb{H}_D(\mathcal{P}^e)\}} \right) \end{array} \right\}. \tag{8}$$

Thus, given the platform being chosen by the R party, the D party sets a platform that achieves the highest social welfare among voters who have a preference for the D party on the basis of $\mathcal{P}^e$. Importantly, the social welfare of a party from a given platform takes account of the likelihood of the party winning the national election on basis of that platform. And, at this choice stage, the parties contemplate platforms that are different from what voters expect. Of course, in equilibrium the chosen platforms will coincide with expected ones.

Analogously, the R party platform solves

$$(w_R, z_R) = \underset{(w,z)}{\operatorname{argmax}} \left\{ \begin{array}{l} [1-\pi(w_D, z_D, w, z)]\mathbb{E}_{(x,y)}\left(\left[-(w-x)^2-(z-y)^2\right]\mathbb{1}_{\{(x,y)\in\mathbb{H}_R(\mathcal{P}^e)\}}\right) \\ +\pi(w_D, z_D, w, z)\,\mathbb{E}_{(x,y)}\left(\left[-(w_D^e-x)^2-(z_D^e-y)^2\right]\mathbb{1}_{\{(x,y)\in\mathbb{H}_R(\mathcal{P}^e)\}}\right) \end{array} \right\}. \quad (9)$$

3.4 Equilibrium

Definition 1. $(w_D^*, z_D^*, w_R^*, z_R^*)$ *is an equilibrium if* $\mathcal{P}^e = (w_D^*, z_D^*, w_R^*, z_R^*)$ *implies* (w_D^*, z_D^*) *solves* (8) *and* (w_R^*, z_R^*) *solves* (9). *The equilibrium is* ***convergent*** *if* $(w_D^*, z_D^*) = (w_R^*, z_R^*)$, *otherwise it is nonconvergent or* ***polarized****.*

Going forward, we will focus on polarized equilibria since this is the case most relevant for us. The case of convergent equilibria is discussed briefly in Appendix B.

4 Polarized Equilibria

Without loss of generality, we will consider polarized equilibria in which the parties differ along the x dimension and may or may not differ along the y dimension. Thus, throughout this section it is assumed that $w_D \neq w_R$ and $w_D^e \neq w_R^e$.

4.1 Characterization of $\mathbb{A}(\mathcal{P})$

Let $\mathcal{P} = (w_D, z_D; w_R, z_R)$ denote the party platforms in the election. Given $\mathcal{P}$ and aggregate shock A, for each $y \in \mathbb{R}$ there exists a threshold $x(y, A, \mathcal{P}) \in \mathbb{R}$ such that a voter of type $(x(y, A, \mathcal{P}), y)$ is indifferent between the two parties. This threshold value is given by

$$x(y, A, \mathcal{P}) = \frac{\left[(z_D-y)^2-(z_R-y)^2\right] - A + \left[w_D^2-w_R^2\right]}{2\left[w_D-w_R\right]}. \quad (10)$$

If $w_D > w_R$, we may verify that $U(x, y; w_D, z_D) + A - U(x, y; w_R, z_R)$ is increasing in x. In this case a voter of type (x, y) votes for the D party if $x \geq \ x(y, A, \mathcal{P})$ and votes for the R party if $x < \ x(y, A, \mathcal{P})$. On the other hand, if $w_D < w_R$ then $U(x, y; w_D, z_D) + A - U(x, y; w_R, z_R)$ is decreasing in x and a voter of type (x, y) votes for the D party if $x \leq \ x(y, A, \mathcal{P})$ and votes for the R party if $x > \ x(y, A, \mathcal{P})$. Regardless, an increase in A expands the set of voters who vote for D party.

Define $\mathbb{A}(\mathcal{P})$ as the set of A values for which the D party gets at least 50 percent of the votes in the general election, given the (combined) party platform $\mathcal{P}$. We can now give our first characterization result.

Proposition 1. *Let $\bar{A}(\mathcal{P})$ be such that $x(0, \bar{A}(\mathcal{P}), \mathcal{P}) = 0$. Then $\bar{A}(\mathcal{P}) = [w_D^2 + z_D^2] - [w_R^2 + z_R^2]$ and $\mathbb{A}(\mathcal{P}) = \{A \in \mathbb{R} : A \geq \bar{A}(\mathcal{P})\}$.*

Proof. $\bar{A}(\mathcal{P})$ to be such that $x(0, \bar{A}(\mathcal{P}), \mathcal{P}) = 0$. Then From (10) we have

$$x(0, A, \mathcal{P}) = \frac{[z_D^2 - z_R^2] - A + [w_D^2 - w_R^2]}{2[w_D - w_R]}.$$

Hence $x(0, \bar{A}(\mathcal{P}), \mathcal{P}) = 0$ implies that $\bar{A}(\mathcal{P}) = [w_D^2 + z_D^2] - [w_R^2 + z_R^2]$. Next, note that

$$x(y, A, \mathcal{P}) = x(0, A, \mathcal{P}) + y \left[\frac{z_R - z_D}{w_D - w_R} \right],$$

from which it follows

$$x(y, \bar{A}(\mathcal{P}), \mathcal{P}) = y \left[\frac{z_R - z_D}{w_D - w_R} \right].$$

Now observe that if for (x, y) we have $x > x(y, \bar{A}(\mathcal{P}), \mathcal{P})$, then for $(-x, -y)$ we must have $-x < x(-y, \bar{A}(\mathcal{P}), \mathcal{P})$. Given (2), the measure of voters voting for each of the two parties must be exactly equal and therefore equal to one-half. Since the l.h.s. of (3) is increasing in A, it follows that D party will achieve at least a majority if $A \geq \bar{A}(\mathcal{P})$. Hence $\mathbb{A}(\mathcal{P}) = \{A \in \mathbb{R} : A \geq \bar{A}(\mathcal{P})\}$. □

4.2 Characterization of $H_k(\mathcal{P}^e)$

Let $\mathcal{P}^e$ denote $(w_D^e, z_D^e, w_R^e, z_R^e)$.

Proposition 2. *For each $y \in \mathbb{R}$ there exists $\bar{x}(y, \mathcal{P}^e)$ such that (i) if $w_D^e > w_R^e$ then $H_D(\mathcal{P}^e) = \{(x,y) \in \mathbb{R}^2 : x \geq \bar{x}(y, \mathcal{P}^e)\}$ and $H_R(\mathcal{P}^e) = \{(x,y) \in \mathbb{R}^2 : x \leq \bar{x}(y, \mathcal{P}^e)\}$ and (ii) if $w_D^e < w^e R$ then $H_D(\mathcal{P}^e) = \{(x,y) \in \mathbb{R}^2 : x \leq \bar{x}(y, \mathcal{P}^e)\}$ and $H_R(\mathcal{P}^e) = \{(x,y) \in \mathbb{R}^2 : x \geq \bar{x}(y, \mathcal{P}^e)\}$.*

Proof. Given y and $\mathcal{P}^e$, define $\bar{x}(y, \mathcal{P}^e)$ to be the value of x for which the voter is indifferent between the party platforms. Then $\bar{x}(y, \mathcal{P}^e)$solves:

$$-(w_D^e - \bar{x}(y, \mathcal{P}^e))^2 - (z_D^e - y)^2 = -(w_R^e - \bar{x}(y, \mathcal{P}^e))^2 - (z_R^e - y)^2, \tag{11}$$

which implies

$$\bar{x}(y, \mathcal{P}^e) = \frac{(z_D^e - y)^2 - (z_R^e - y)^2 + [w_D^{e^2} - w_R^{e^2}]}{2(w_D^e - w_R^e)}. \tag{12}$$

If $w_D^e > w_R^e$, the expression $-(w_D^e - x)^2 + (w_R^e - x)^2$ is strictly increasing in x. Then, given y, for all $x \geq (\leq)\ \bar{x}(y, \mathcal{P}^e)$, the l.h.s. of (11) is at least as large as (not greater than) the r.h.s. and $H_D(\mathcal{P}^e) = \{(x,y) \in \mathbb{R}^2 : x \geq \bar{x}(y, \mathcal{P}^e)\}$ and $H_R(\mathcal{P}^e) = \{(x,y) \in \mathbb{R}^2 : x \leq \bar{x}(y, \mathcal{P}^e)\}$. If $w_D^e < w_R^e$, then the opposite implications follow and $H_D(\mathcal{P}^e) = \{(x,y) \in \mathbb{R}^2 : x \leq \bar{x}(y, \mathcal{P}^e)\}$ and $H_R(\mathcal{P}^e) = \{(x,y) \in \mathbb{R}^2 : x \geq \bar{x}(y, \mathcal{P}^e)\}$. $\square$

There is clear geometric intuition for the expression for $\bar{x}(y, \mathcal{P}^e)$. Expanding the squared terms in (12), $\bar{x}(y, \mathcal{P}^e)$ can be expressed as:

$$\bar{x} = \frac{{w_D^e}^2 + {z_D^e}^2 - [{w_R^e}^2 + {z_R^e}^2]}{2(w_D^e - w_R^e)} - \left[\frac{z_D^e - z_R^e}{w_D^e - w_R^e}\right] y. \tag{13}$$

This is a straight line passing through the midpoint of the line joining (w_D^e, z_D^e) and (w_R^e, z_R^e) and orthogonal to it. All points on this line are equidistant from (w_D^e, z_D^e) and (w_R^e, z_R^e) and, therefore, voter types with ideal points on this line get the same utility from either platform (since utility *is* (negative of) the Euclidean distance from (w_k^e, z_k^e)). Thus, voters' beliefs about party platforms divide the issue space into two closed half-planes, with voter types in the interior of these half-planes having strict preference for one of the parties.

4.3 Characterizing the Best Responses of Parties

The (common) beliefs of voters determine which primary election they participate in. The parties take these primary participation decisions as given and choose their policy platform strategically, i.e., taking into account the policy platform of the other party.

The D-party's choice problem given in (8) can be expressed as

$$\max_{(w,z)} \left\{ \begin{array}{c} -L(w_R, z_R) + \\ \left[1 - F\left(w^2 + z^2 - [{w_R}^2 + {z_R}^2]\right)\right] \times \\ m(H_D(\mathcal{P}^e))\mathbb{E}_{(x,y)|H_D(\mathcal{P}^e)} \left\{- \left[(w-x)^2 + (z-y)^2\right] + \left[(w_R - x)^2 + (z_R - y)^2\right]\right\} \end{array} \right\}, \tag{14}$$

where $m(H_D(\mathcal{P}^e))$ is the measure of the set people who attend the D-party primary, $\mathbb{E}$ is expectation taken with respect to the distribution of (x, y) conditional on $(x, y) \in H_D(\mathcal{P}^e)$, and $L(w_R, z_R)$ is shorthand for $\mathbb{E}_{(x,y)|H_D(\mathcal{P}^e)}[(w_R - x)^2 + (z_R - y)^2]$. Ignoring the constant term L, the objective function is thus the product of the probability of D party winning the national election and the total gain of the people attending the primary from enjoying D party's policies rather than those of the R party.

Analogously, but recognizing that R party's probability of winning the election is $F(\bar{A}(\mathcal{P}))$, R party's choice problem reduces to

$$\max_{(w,z)} \left\{ \begin{array}{c} -L(w_D, z_D) + \\ F\left(w_D^2 + z_D^2 - [w^2 + z^2]\right) \times \\ m(H_R(\mathcal{P}^e))\mathbb{E}_{(x,y)|H_R(\mathcal{P}^e)} \left\{- \left[(w-x)^2 + (z-y)^2\right] + \left[(w_D - x)^2 + (z_D - y)^2\right] - L(W_D, z_D)\right\} \end{array} \right\}. \tag{15}$$

If the k party cared *only* about winning the national election, its optimal policy would be to set $(w_k, z_k) = (0, 0)$ because this choice maximizes the probability of a k-party win. On the other hand, if it cared *only* about maximizing the total gain conditional on a win, its optimal policy would be set to $(w_k, z_k) = (\mathbb{E}(x|H_k(\mathcal{P}^e)), \mathbb{E}(y|H_k(\mathcal{P}^e))$ as $\mathbb{E}_{(x,y)|H^k(\mathcal{P}^e)}\left[(w-x)^2 + (z-y)^2\right]$ is minimized if w_k and z_k are set to the relevant conditional means of x and y, respectively.

Generally speaking, these two goals are mutually incompatible and the party's best response strikes a balance between them.[5] For the case where at least one conditional mean is nonzero, we rely on first-order conditions to characterize a party's best response. Assume that $m_k(\mathcal{P}^e) > 0$ so that all the conditional expectation terms are well-defined. If at the point of best response $F'(\cdot)$ exists and $1 - F(\cdot) > 0$, the best responses must satisfy the following first-order conditions:

$$(1+\phi_k)w = \mathbb{E}_{(x,y)|H_k(\mathcal{P}^e)}x \quad \text{and} \quad (1+\phi_k)z = \mathbb{E}_{(x,y)|H_k(\mathcal{P}^e)}y, \tag{16}$$

where

$$\phi_D = \left[\frac{F'\left(w^2+z^2-w_R^2-z_R^2\right)}{1-F\left(w^2+z^2-w_R^2-z_R^2\right)}\right] \times \\ \mathbb{E}_{(x,y)|H_D(\mathcal{P}^e)}\{-(w-x)^2-(z-y)^2+[(w_R-x)^2+(z_R-y)^2]\}, \tag{17}$$

and

$$\phi_R = \left[\frac{F'\left(w_D^2+z_D^2-w^2-z^2\right)}{1-F\left(w_D^2+z_D^2-w^2-z^2\right)}\right] \times \\ \mathbb{E}_{(x,y)|H_R(\mathcal{P}^e)}\{-(w-x)^2-(z-y)^2+[(w_D-x)^2+(z_D-y)^2]\}. \tag{18}$$

Note that ϕ_D and ϕ_R can be negative only if the expectation terms in (17) and (18), respectively, are negative. However, this would imply that the optimal utility is lower than what the k party could get by replicating the policies of the $\sim k$ party. Therefore, at an optimum, $\phi \geq 0$. Thus, the optimal (w, z) for each party is on the ray connecting $(0, 0)$ to $(\mathbb{E}_{(x,y)|H_k(\mathcal{P}^e)}x, \mathbb{E}_{(x,y)|H_k(\mathcal{P}^e)}y)$. This shows clearly that parties balance the two polar objectives mentioned earlier (maximizing the gain conditional on a win and maximizing the probability of a win).

The first-order condition also shows that the extent to which w and z move toward the origin and away from their respective conditional means depends in an intuitive way on the distribution of A. The more concentrated the distribution, i.e., the larger F' is, the bigger is ϕ_k and closer to the origin is k party's optimal choice: a party is willing to put up with a platform closer to the origin if moving towards its mean preferences lowers the probability of a win substantially.

[5]In Roemer (2001), these two extremes figure as motivations ascribed to "opportunists" (who care only about winning) and "militants" (who care only about the gain, conditional on winning). What we are referring to as the social planner is called "reformists".

Reliance on first-order conditions requires that they also be sufficient. It is challenging to ensure sufficiency for the general case, but it can be ensured if the distribution of A and the opposing party's platform are jointly restricted.

Proposition 3. *Suppose* $(\mathbb{E}(x|H_k(\mathcal{P}^e)), \mathbb{E}(y|H_k(\mathcal{P}^e))) \neq (0,0)$ *and let* (w_k^*, z_k^*) *be party k's best response to* $(w_{\sim k}, z_{\sim k})$*. Assume that A is uniformly distributed over the support* $[-\alpha, \alpha]$*,* $\alpha > 0$*, and that party k (weakly) prefers the centrist platform* $(0,0)$ *to* $(w_{\sim k}, z_{\sim k})$*, i.e.,* $\mathbb{E}_{(x,y)|H_k(\mathcal{P}^e)}\{-x^2 - y^2 + [(w_{\sim k} - x)^2 + (z_{\sim k} - y)^2]\} \geq 0$*. Let* $C_D \subset \mathbb{R}^2$ *be the set of all* (w, z) *such that all the party D platforms for which neither victory nor defeat is certain in the national election given* (w_R, z_R)*, i.e., all* (w, z) *such that* $-\alpha < w^2 + z^2 - [w_R^2 + z_R^2] < \alpha$*. Similarly, let* $C_R \subset \mathbb{R}^2$ *be the set of all* (w, z) *such that* $-\alpha < w_D^2 + z_D^2 - [w^2 + z^2] < \alpha$*. If* $(\hat{w}_k, \hat{z}_k) \in C_k$ *solves the corresponding first order condition* (16) *given* $(w_{\sim k}, z_{\sim k})$*, then* $(\hat{w}_k, \hat{z}_k) = (w_k^*, z_k^*)$*.*

Proof. See Appendix A

Two remarks: First, the requirement that the first-order conditions have solutions in C_k is important: At the putative optimum, neither victory nor defeat should be certain. For instance, if the best response is such that a victory is certain, the first-order conditions will not be satisfied. This is because such an optimum is always located at a kink in the objective function: $F'(\cdot)$ is zero whenever the change in policy brings the policy closer to 0 and it is $-1/[2\alpha]$ whenever the change takes the policy further way from zero. Second, the condition $\mathbb{E}_{(x,y)|H_k(\mathcal{P}^e)}\{-x^2 - y^2 + [(w_{\sim k} - x)^2 + (z_{\sim k} - y)^2]\} \geq 0$ is key to the objective function being concave over the relevant domain so that first-order conditions are both necessary and sufficient for optimality.

4.4 Symmetric (Polarized) Equilibria

To characterize the equilibrium further, we will focus on symmetric equilibria which seems natural given the environmental symmetry of our model. A (polarized) equilibrium is **symmetric** if $(w_D^*, z_D^*) = -(w_R^*, z_R^*)$. In addition, we will assume that the distribution of voter ideal points is uniform on a circle centered at $(0,0)$ with radius $\theta > 0$.

The determination of symmetric equilibria can be conveniently broken down into three stages. In the first stage, we take the mean preferences of the voters attending each primary as given and solve for the Nash equilibrium of the policy game between the two parties using the general results of the previous section. In the second stage, we solve for the mean preferences of the voters attending each primary, given common beliefs about party platforms. In the final part, we combine the first two parts to determine the common equilibrium beliefs, i.e., beliefs that generate policy platforms that produce those beliefs.

Conditional Means Determine Party Platforms: We focus on the case where the mean preferences of people attending the two primaries are symmetric to each other, i.e., $(\mathbb{E}_D x, \mathbb{E}_D y) = (-\mathbb{E}_R x, -\mathbb{E}_R y)$. As will be shown in the next (sub)section, this symmetry is implied whenever the common beliefs about party platforms are symmetric (as they must be, by definition, in a symmetric equilibrium).

Proposition 4. *Let $A \sim U[-\alpha, \alpha]$. Given $(\mathbb{E}_D x, \mathbb{E}_D y) = -(\mathbb{E}_R x, \mathbb{E}_R y) \neq (0,0)$, the symmetric Nash equilibrium of the policy game has $(w_k^*, z_k^*) = (h^* \mathbb{E}_k x, h^* \mathbb{E}_k y)$, $k \in \{D, R\}$, where $h^* \in (0,1)$ is given by*

$$h^* = \frac{-\alpha + \sqrt{16\alpha[(\mathbb{E}_k x)^2 + (\mathbb{E}_k y)^2] + \alpha^2}}{8[(\mathbb{E}_k x)^2 + (\mathbb{E}_k y)^2]}, \ k \in \{D, R\}. \tag{19}$$

Proof. In a symmetric equilibrium, $(w_k^*, z_k^*) = (-w_{\sim k}^*, -z_{\sim k}^*)$. Making this substitution in the k party's first-order condition (16), using the fact that $F'(0) = 1/[2\alpha]$ and $F(0) = 1/2$, and substituting $h^* \mathbb{E}_k x$ and $h^* \mathbb{E}_k y$ for w_k^* and z_k^* respectively, lead to the following pair of first-order conditions:

$$[\alpha + 4h^*(\mathbb{E}_k x)^2 + 4h^*(\mathbb{E}_k x)^2] h^* \mathbb{E}_k x = \alpha \mathbb{E}_k x \tag{20}$$

$$[\alpha + 4h^*(\mathbb{E}_k x)^2 + 4h^*(\mathbb{E}_k x)^2] h^* \mathbb{E}_k y = \alpha \mathbb{E}_k y. \tag{21}$$

Since at least one of the conditional means is non-zero, eliminating the non-zero conditional mean from the appropriate first-order condition leads to a quadratic equation in h^*, and equation (19) is the positive solution of this quadratic (since h^* is just $(1+\phi)^{-1}$ and $\phi \geq 0$ at an optimum, the positive solution is the relevant one). We can easily verify that $h^* < 1$.

Since we used first-order conditions to solve for h^* we must verify that the first-order conditions are sufficient for optimality. For party k, $(w_{\sim k}, z_{\sim k}) = (-w_k^*, -z_k^*)$. Hence $C_k = \{(w,z) : -\alpha < w^2 + z^2 - [(-w_k^*)^2 + (-z_k^*)^2] < \alpha\}$ and clearly (w_k^*, z_k^*) is in C_k (since $-\alpha < 0 < \alpha$). Next, we may verify that the average gain for party k from the centrist platform is $(h^* \mathbb{E}_k x)^2 + (h^* \mathbb{E}_k y)^2 +$

$2[h^*(\mathbb{E}_k x)^2 + h^*(\mathbb{E}_k y)^2] > 0$. The inequality follows since at least one conditional mean is non-zero and $h^* > 0$.

$\square$

Beliefs Determine Conditional Means: We consider all $\mathcal{P}^e$ such that $(w_D^e, z_D^e) = -(w_R^e, z_R^e)$ and $w_D^e \neq 0$. Thus, beliefs are symmetrically polarized along the x dimension and may or may not be (symmetrically) polarized along the y dimension.

To go from (symmetric) beliefs to conditional means requires us to be specific about the distribution of voter ideal points. As noted earlier, the distribution of voter ideal points is assumed to be uniform on a circle centered at $(0, 0)$ with radius $\theta > 0$. The density of voter ideal points is then given by

$$f(x, y) = \begin{cases} \frac{1}{\pi\theta^2} & \text{if } 0 \leq x^2 + y^2 \leq \theta \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

Clearly, $f(x, y)$ is symmetric; it is also (trivially) *radially* symmetric meaning that any two points that are equally distant from the center have the same density.

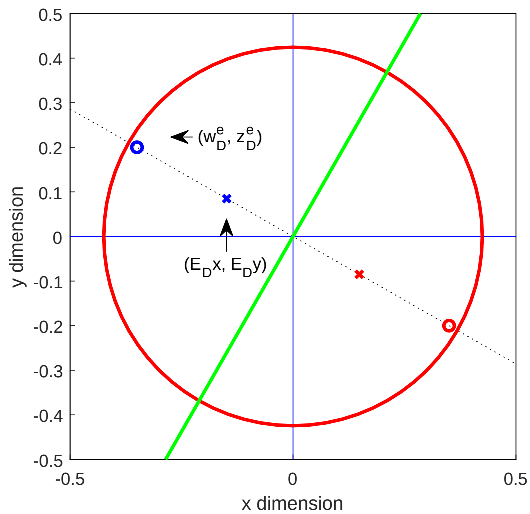
Given $w_D^e \neq 0$ and symmetry of beliefs, it follows from (13) that the line separating D-party primary attendees from R-party primary attendees is given by

$$\bar{x}(y; \mathcal{P}^e) = -\left[\frac{z_D^e - z_R^e}{w_D^e - w_R^e}\right] y = -\left[\frac{z_D^e}{w_D^e}\right] y.$$

This is a straight line that goes through the origin and divides the circular support of the voters ideal point distribution into two semi-circles.

Figure 4 shows an example where points marked by circles represent the common symmetric beliefs about the platforms of the D and R parties, respectively. In this figure, $w_D^e < 0$ and $z_D^e > 0$ and, so, the line separating voters is the positively sloped straight line going through the origin and perpendicular to the dotted line joining the circle points. All voters with ideal points that lie in the semi-circle to the left of the straight line attend the D-party primary and all voters with ideal points in the semi-circle to the right of the straight line attend the R-party primary.

Figure 4:
Partisan Sorting Based on Beliefs



Since the distribution of the voter ideal points is uniform over each semi-circle, the points $(\mathbb{E}_D x, \mathbb{E}_D y)$ and $(\mathbb{E}_R x, \mathbb{E}_R y)$ coincide with the center of gravity, or centroid, of the D-party and R-party semi-circles, respectively. This allows for an easy computation of conditional means $(\mathbb{E}_k x, \mathbb{E}_k y)$ using the fact that the centroid of a semi-circle of radius θ lies on the radius that is perpendicular to the base of the semi-circle and a distance $4\theta/3\pi$ from the center. In Figure 4, the conditional means of the D- and R-party semi-circles are denoted by points marked by 'x'.

With these facts in mind, we have the following:

Proposition 5. *Let $f(x, y)$ be given by (22). Let $\mathcal{P}^e$ be a symmetric belief vector, i.e., $(w^e_D, z^e_D) = -(w^e_R, z^e_R) \neq (0, 0)$. For concreteness, assume $w^e_D \neq 0$ and let $\beta = z^e_k/w^e_k$. Then,*

$$\begin{bmatrix} \mathbb{E}_k x \\ \mathbb{E}_k y \end{bmatrix} = \begin{bmatrix} \mathrm{sgn}(w^e_k)\frac{4\theta}{3\pi}\frac{1}{\sqrt{1+\beta^2}} \\ \mathrm{sgn}(z^e_k)\frac{4\theta}{3\pi}\sqrt{\frac{\beta^2}{1+\beta^2}} \end{bmatrix}. \tag{23}$$

Proof. Since the radius of the semi-circles shown in Figure 4 is θ, it follows that $(\mathbb{E}_k x)^2 + (\mathbb{E}_k y)^2 = \left(\frac{4\theta}{3\pi}\right)^2$. Next, observe that since the conditional mean is located on the radius perpendicular to the base, $\mathbb{E}_k y/\mathbb{E}_k x = w^e_k/z^e_k = \beta$. Using these two facts we get

$$(\mathbb{E}_k x)^2 = \left[\frac{4\theta}{3\pi}\right]^2 \frac{1}{1+\beta^2} \text{ and } (\mathbb{E}_k y)^2 = \beta^2(\mathbb{E}_k x)^2. \tag{24}$$

These equations determine the magnitudes of $\mathbb{E}_k x$ and $\mathbb{E}_k y$ but not their signs. To determine their signs, note that $(\mathbb{E}_k x, \mathbb{E}_k y)$ is on the line connecting $(0,0)$ to (w_k^e, z_k^e) and, so, the signs of $\mathbb{E}_k x$ and $\mathbb{E}_k y$ must be the signs of w_k^e and z_k^e, respectively. Combining these facts leads to equation (23).

$\square$

Equilibrium: For a (symmetric) belief $\mathcal{P}^e$ to be an equilibrium belief, the conditional means implied by these beliefs must generate policy outcomes that validate those beliefs. The radial symmetry of the ideal point distribution implies that there is a continuum of symmetric equilibria indexed by the direction of the (w_D^e, z_D^e) vector, with *any* direction being consistent with an equilibrium. Specifically, we can construct a symmetric equilibrium by performing the following steps:

- Pick a $\beta^2 \geq 0$ and pick signs for w_D^e and z_D^e — these choices fix the (w_D^e, z_D^e) vector and by symmetry also fix the (w_R^e, z_R^e) vector.
- Use these settings in (23) to determine $(\mathbb{E}_k x, \mathbb{E}_k y)$ for $k \in \{D, R\}$.
- Use the resulting conditional means in (19) to determine h^*.
- Then $\mathcal{P}^{e*} = (h^* \mathbb{E}_D x, h^* \mathbb{E}_D y, h^* \mathbb{E}_R x, h^* \mathbb{E}_R y)$ is an equilibrium belief.

This construction works for two reasons. First, for symmetric beliefs, the outcome of the policy game is determined by the *direction* of beliefs, i.e., by β^2 and the signs of w_D^e and z_D^e, and not the values of (w_k^e, z_k^e).[6] Second, the radial symmetry of the uniform circular distribution implies that the direction of the conditional mean vector always coincides with the direction of the belief vector — this is the property that $(\mathbb{E}_D x, \mathbb{E}_D y)$ is located on the ray connecting to $(0,0)$ to (w_D^e, z_D^e) (similarly for the conditional mean of the R-party). Since the equilibrium platform of the D party lies on the ray connecting $(0,0)$ to $(\mathbb{E}_D x, \mathbb{E}_D y)$ (similarly for the equilibrium platform of the R party), the equilibrium platform is also on the ray connecting $(0,0)$ to (w_D^e, z_D^e). This means that we can choose a direction for the belief vector and then pick the *position* of the belief vector to match outcomes. Put slightly differently, we can pick any symmetric belief vector to start off, compute the direction of the belief vectors, compute the equilibrium platforms that result from

[6] This is intuitive. Scaling the belief vector up or down does not change β and, so, does not change the sorting of the voters into the two primaries and therefore does not change the outcome of the policy game.

these directions, then simply scale the initial belief vector to match the equilibrium platforms. We summarize this in the following proposition.

Proposition 6. *Let $\mathcal{P}^e = (w_D^e, z_D^e, w_R^e, z_R^e)$ be any symmetric belief vector with $w_D^e \neq 0$. Then, there exists a constant $\lambda^* > 0$ such that $(\lambda^* w_D^e, \lambda^* z_D^e, \lambda^* w_R^e, \lambda^* z_R^e)$ is an equilibrium belief vector.*

The existence of a continuum of symmetric equilibria reflects the result of sorting of voters into parties. If voters come to believe that party D stands for certain types of policies, then voters who prefer those types of policies will come to determine the composition of the party and, so, its policies.

We close this section with the key equilibrium implication of the model, namely, across equilibria there is an inverse association between the ideological distances between the two parties along x and y dimensions. Define the ideological distance between the two parties along x dimension as $\delta_x = |w_D^e - w_R^e|$ and along the y dimension as $\delta_y = |z_D^e - z_R^e|$. In a symmetric equilibrium, $\delta_x = 2|w_D^e|$ and $\delta_y = 2|z_D^e|$.

Proposition 7. *Let $\mathcal{P}^e$ and $\mathcal{P}'^e$ be any two symmetric equilibrium belief vectors, with w_D^e and $w_D'^e$ not equal to zero. Then, $\delta_x > \delta_x'$ if and only if $\delta_y < \delta_y'$.*

Proof. Let $\beta = z_D^e / w_D^e$ and $\beta' = z_D'^e / w_D'^e$. By Proposition 5 and 6, we have that

$$|w_D^e| = h^* \left[\frac{4\theta}{3\pi}\right] \frac{1}{\sqrt{1+\beta^2}} \text{ and } |w_D'^e| = h'^* \left[\frac{4\theta}{3\pi}\right] \frac{1}{\sqrt{1+\beta'^2}}, \tag{25}$$

$$|z_D^e| = h^* \left[\frac{4\theta}{3\pi}\right] \sqrt{\frac{\beta^2}{1+\beta^2}} \text{ and } |z_D'^e| = h'^* \left[\frac{4\theta}{3\pi}\right] \sqrt{\frac{\beta'^2}{1+\beta'^2}}. \tag{26}$$

By equations (19) and (24), we have that

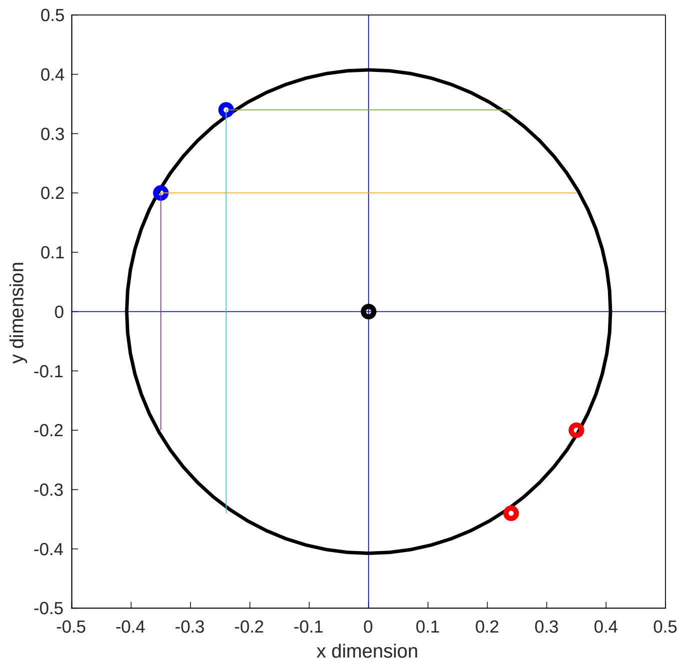
$$h^* = \frac{-\alpha + \sqrt{16\alpha \left[\frac{4\theta}{3\pi}\right]^2 + \alpha^2}}{8\left[\frac{4\theta}{3\pi}\right]^2} = h'^*. \tag{27}$$

Therefore, if $\delta_x > \delta_x'$ then $\beta < \beta'$ and so $\delta_y' < \delta_y$. Conversely, if $\delta_y' < \delta_y$ then $\beta < \beta'$ and, so, $\delta_x > \delta_x'$. □

The intuition for this result is quite simple. In our model, equilibrium policy platforms (equivalently, equilibrium party ideologies) are a pair of diametrically opposite points on a circle of radius

$h^* 4\theta/3\pi$ centered at $(0, 0)$ and any such pair of points is an equilibrium as shown in Figure 5. As we move along the circle, we trace out different equilibria. Comparing across these equilibria, the ideological distance along the x dimension can increase only at the expense of the ideological distance along the y dimension and vice versa.

Figure 5:
Partisan Sorting Based on Beliefs

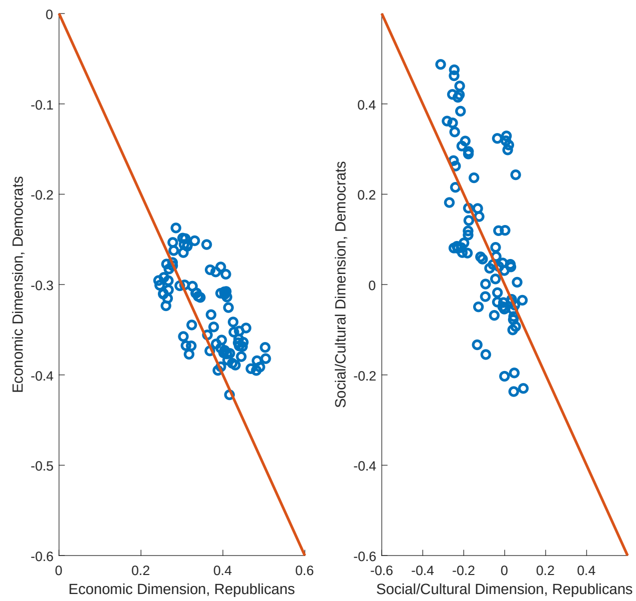


5 Secular Changes in Polarization of Political Parties in the US

Up to this point, we have presented a model in which polarization along one dimension interacts with polarization along the other dimension. The interaction arises through the value of β^2. To recall, as β^2 increases, polarization along the x dimension decreases and that along the y dimension increases. In this section, we ask if this implication of the model has quantitative bite. That is, if the model is supplied with an empirical analog of β^2 for each Congress, can it reproduce the ideological distances for each Congress we see in Figures 1 and 2?

We are motivated to do this because historical experience seems reasonably consonant with a key maintained assumption of the theory, namely, that policy outcomes are symmetric. Figure 6 gives scatter plots of the ideological positions of the two parties for the two dimensions. In each

Figure 6:
Approximate Symmetry of Ideological Positions



Source: https://voteview.com/ and authors' calculations

panel, the solid downward sloping line has a slope of -1 and passes through the origin $(0, 0)$. If the ideological positions of the two parties for a dimension were exactly symmetrically opposed, all points would lie on the solid line. While not exactly symmetric, the actual points cluster around the (negative) 45 degree line. Furthermore, the fact that ideological positions of the two parties have varied considerably while remaining roughly symmetric suggests that the assumption of a radially symmetric distribution of voter ideal points is not too far off the mark.

For a given Congress, we take the empirical analog of β^2, denoted $\hat{\beta}^2$, to be $([D2 - R2]/[D1 - R1])^2$, where $D1$ and $D2$ are the mean ideological position of Democrats for the liberal-conservative and social/cultural dimensions, respectively. In the theory, these correspond to w_D and z_D, respectively. Similarly, $R1$ and $R2$ are the mean ideological positions for the liberal-conservative and social/cultural dimensions for the Republicans and w_R and z_R are the theory counterparts. The theory counterpart of $\hat{\beta}^2$ is β^2 since $w_R = -w_R$ and $z_D = -z_R$ and $z_D/w_D = \beta$.

In order to connect $\hat{\beta}^2$ to the theoretically predicted ideological distances, we need values of θ and α. For this, note that in any equilibrium the (common) Euclidean distance of a party's platform from the origin is a constant given by

$$
\begin{aligned}
\sqrt{(w_k^*)^2 + (z_k^*)^2} =& \sqrt{h^{*2}[(\mathbb{E}_k x)^2 + (\mathbb{E}_k y)^2]} \\
=& h^* \left[\frac{4\theta}{3\pi}\right] \\
=& \frac{-\alpha + \sqrt{16\alpha\left[\frac{4\theta}{3\pi}\right]^2 + \alpha^2}}{8\left[\frac{4\theta}{3\pi}\right]}.
\end{aligned} \tag{28}
$$

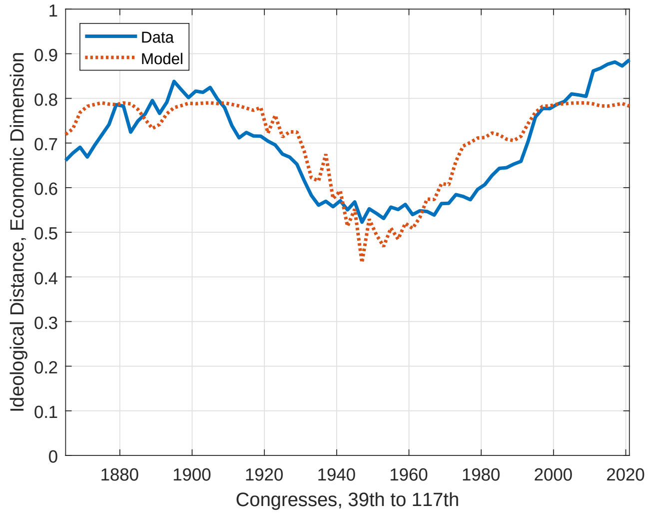
In actuality, the Euclidean distance (from the origin) of a party's ideology has not been constant over time and neither have the Euclidean distances of the two parties been exactly equal to each other at all times. However, it is the case that the distances of the D-party platform averaged over time is quite close to the distances of the R-party platform averaged over time, these being 0.39 and 0.40, respectively. This closeness of the time averages is another reflection of the rough symmetry of ideological positions of the two parties. We set $\theta = 1$ and pick α so that the expression in (28) is equal to $(0.39 + 0.40)/2$.[7] With these settings of θ and α, we use time series on $\hat{\beta}^2$ to get the theoretically predicted time series for δ_x and δ_y using the expressions in (25) and (26).

Figures 7 and 8 plots δ_x along with $|D1 - R1|$ and δ_y along with $|D2 - R2|$, respectively. In these figures, the solid line is history and the dotted line is the prediction of the calibrated model. In both figures, the predicted ideological distance tracks the actual ideological distances, with the correspondence being remarkably close for δ_y and $|D2 - R2|$. These plots reveal that knowledge of the *ratio* $[(D2 - R2)/(D1 - R1)]^2$, which is $\hat{\beta}^2$, allows prediction of the *levels* of $|D2 - R2|$ and $|D1 - R1|$. Thus, during the '40s and '50s, when $\hat{\beta}^2$ was rising, the model predicts that $|D1 - R1|$ must fall as observed, and in the post-1965 era when $\hat{\beta}^2$ fell, the model predicts that $|D1 - R1|$ should rise, as observed.

The premise on which this effect is based is diversity with respect to the social/cultural ideology among both economic liberals and economic conservatives. When a polity is polarized along the social dimension, it is possible — i.e., it can be an equilibrium outcome — that modest economic conservatives (liberals) who are socially very conservative (liberal) can align themselves with the

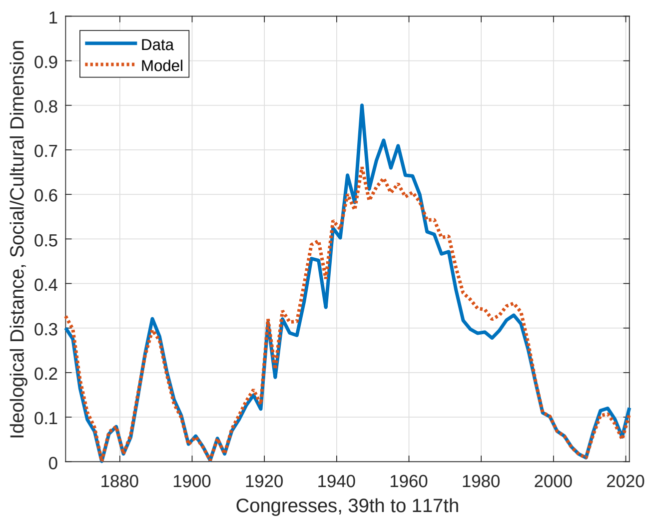
[7] If we had picked a different value for θ but altered the value of α so that the expression in (28) evaluated to 0.395, the theoretically predicted equilibrium outcomes would remain unchanged. In this sense, the choice of θ is a normalization. We note, however, that the method employed by the DW-Nominate program to infer the ideological positions of members of Congress assumes that the position is always contained in $[-1, 1] \times [-1, 1]$. Thus, by construction, the mean ideological position of a party along any dimension cannot exceed 1 in absolute value. For our model to be consistent with this restriction, θ cannot exceed $3\pi/4 = 2.37$; if it did, there will exist values of α and β^2 for which the equilibrium ideological position of parties along at least one dimension will exceed 1.

Figure 7:
Ideological Distance, Economic Dimension
Data and Model Predictions



Source: https://voteview.com/ and authors' calculations

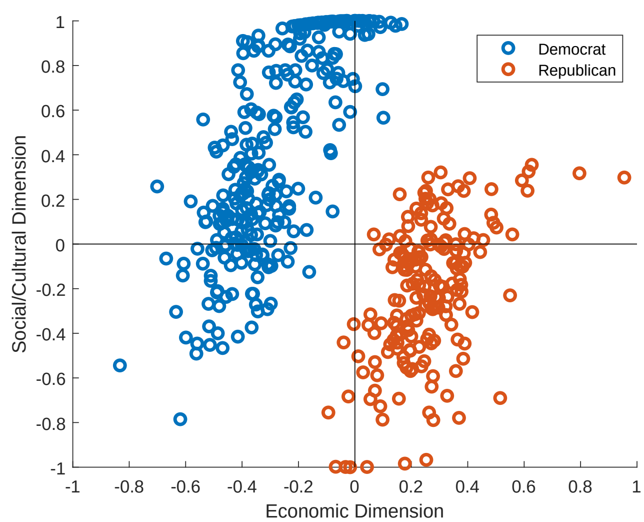
Figure 8:
Ideological Distance, Social/Cultural Dimension
Data and Model Predictions



Source: https://voteview.com/ and authors' calculations economically *liberal* (*conservative*) party because the economically liberal (conservative) party happens also to be the more socially conservative (liberal) party. When they do, the alignment leads to *less* sorting, i.e., *more* homogeneity, along the economic dimension across the two parties and, therefore, smaller ideological distance between the two parties along the economic dimension.

As an example of this type of sorting, Figure 9 displays the ideological positions of members of the House of Representatives for 88th (1963-1965) Congress. Note the presence of economic conservatives in the Democratic party and economic liberals in the Republican party. As a contrast to this type of sorting, Figure 10 displays the ideological map of the 117th (2021-2023) Congress. We no longer see any economic conservatives in the Democratic party or economic liberals in the Republican party. The result of this sorting pattern is that the two parties are more ideologically distant along the economic dimension.

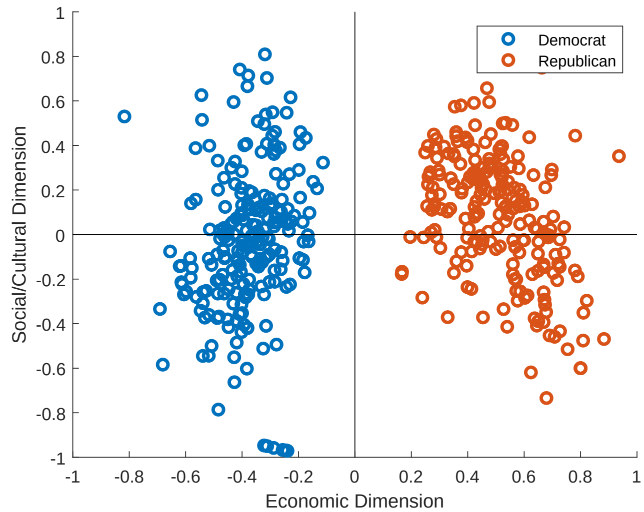
Figure 9:
Member Ideologies, 88th Congress (1963-1965)



Source: https://voteview.com/

These findings bring a novel perspective on the decline and subsequent rise in polarization along the economic dimension in the post-WWII era. In the years leading up to the Civil Rights Act of 1964, the ideological distance between the Democratic and Republican parties widened along the social/cultural dimension. According to our theory, this widening was a key factor in accounting for the *narrowing* of ideological distance along the economic dimension during the same period.

Figure 10:
Member Ideologies, 117th Congress (2021-2023)



Source: https://voteview.com/

Similarly, the narrowing of ideological distance along the social/cultural dimension is an important ingredient in the *widening* of ideological distance along the economic dimension since 1965.

Finally, we note that the increase in polarization that has occurred since 2000 along the economic dimension is not explained by changes in the ideological distance along the social/cultural dimension. Since the turn of the 21st century, changes in the ideological distance along the social/cultural dimension have been small and these changes do not imply significant changes to polarization along the economic dimension, which is predicted to be essentially flat. Thus our results are not in conflict with studies (cited earlier) that have focused on reasons (inequality, campaign contributions, biased beliefs) to account for increasing polarization along the economic dimension in recent decades.

6 Summary

Viewed over a long stretch of time, roll call voting records for Democratic and Republican party representatives in Congress show an inverse association between the mean ideological distance along the social/cultural dimension and the mean ideological distance along the economic dimension.

We presented a model of party competition with a two-dimensional ideology space to explain this fact. In our model, voters have (common) beliefs about the ideological stance of the two parties and join a party based on these beliefs. Each party chooses its national platform by balancing the preferences of the voters who compose the party and the preferences of the polity at large since a party can come to power only if it wins the general election.

If the distribution of voter ideologies is radially symmetric, there is a continuum of polarized equilibria indexed by voter beliefs. Moving from an equilibrium in which the two parties are more polarized along the social/cultural dimension to one in which they are less polarized along this dimension implies an *opposite* movement for polarization along the economic dimension, i.e., from a less economically polarized equilibrium to a more economically polarized equilibrium. This occurs as a result of the differential sorting into the two parties induced by the different beliefs about what each party stands for.

Remarkably, the sorting and re-sorting of a stable distribution of voters into the two parties gives a very good *quantitative* account of the variation in ideological distances along the two dimensions over a long stretch of history. This finding brings a new perspective on the potential causes of political polarization that highlights the role of changes in beliefs as opposed to changes in preferences, law or technology.

References

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

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

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

Azzimonti, M., and M. Fernandes (2021): "Social Media, Networks, Fake News and Polarization," NBER Working Paper No. 24462.

Barber, M., and N. McCarty (2013): "Causes and Consequences of Polarization," in *Negotiating Agreement in Politics, Task Force Report*. American Political Science Association.

Battaglini, M., and S. Coate (2008): "A Dynamic Theory of Public Spending, Taxation and Debt," *American Economic Review*, 98(1), 201–236.

Chatterjee, S., and B. Eyigungor (2020): "Policy Inertia, Election Uncertainty and Incumbency Disadvantage of Political Parties," *Review of Economic Studies*, 87(6), 2600–2638.

Cuadra, G., and H. Sapriza (2008): "Sovereign Defaults, Interest Rates and Political Uncertainty in Emerging Markets," *Journal of International Economics*, 76, 77–88.

De Donder, P., and M. Gallego (2017): "Electoral Competition and Party Positioning," Discussion paper, 17-760, Toulouse School of Economics.

Dixit, A. K., and J. W. Weibull (2007): "Political Polarization," *Proceedings of the National Academy of Sciences*, 104(18), 7351–7356.

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

Drautzburg, T., I. Livshits, and M. L. Wright (2021): "Polarized Contributions but Convergent Agendas," Mimeo.

Herrera, H., D. K. Levine, and C. Martinelli (2008): "Policy Platforms, Campaign Spending and Voter Participation," *Journal of Public Economics*, 92, 501–513.

Konishi, H., and C.-Y. Pan (2020): "Silent Promotion of Agendas: Campaign Contributions and Ideological Polarization," *Public Choice*, 182, 93–117.

Krasa, S., and M. Polborn (2014): "Social Ideology and Taxes in a Differentiated Candidates Framework," *American Economic Review*, 104(1), 308–322.

Lewis, J. B., K. Poole, H. Rosenthal, A. Boche, A. Rudkin, and L. Sonnet (2022): "Voteview: Congressional Roll-Call Votes Database. https://voteview.com," .

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

McCarty, N., K. T. Poole, and H. Rosenthal (2006): *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge: MIT Press.

Poole, K. T. (2005): *Spatial Model of Parliamentary Voting*, Analytical Methods for Social Research. Cambridge University Press.

Poole, K. T., and H. Rosenthal (1997): *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.

Roemer, J. E. (2001): *Political Competition, Theory and Applications*. Harvard University Press.

Wittman, D. (1973): "Parties as Utility Maximizers," *American Political Science Review*, 67, 490–498.

Appendices

A Proof of Proposition 3

We will prove this for $k = D$. For compactness of notation, we will denote $\mathbb{E}(\cdot|H_D(\mathcal{P}^e))$ by use $\mathbb{E}_D(\cdot)$. And, without loss of generality, we will suppose that $\mathbb{E}_D x \neq 0$.

We will show first that $(\hat{w}_D, \hat{z}_D)$ strictly dominates any other point in C_D. Since $(\hat{w}_D, \hat{z}_D)$ solves (16), we have $\hat{z}_D = \beta_D \hat{w}_D$, where $\beta_D = \mathbb{E}_D y / \mathbb{E}_D x$. Within C_D we will only consider alternative (w, z) pairs that also satisfy $z = \beta_D w$. The reason is that for any point in C_D that does not satisfy this equality, there is another point in an ϵ neighbor of it that is also in C_D and that dominates it. Thus all points in C_D where $z \neq \beta_D w$ can be ignored.

Making the substitution for z in the objective function (14) and passing the expectations operator through (and ignoring the constant term) yields:

$$\begin{aligned} G(w) = {} & [1 - F([1+\beta_D^2]w^2 - [w_R^2 + z_R^2])] \times \\ & \{-[1+\beta_D^2]w^2 + 2\mathbb{E}_D x[1+\beta_D^2]\, w - \mathbb{E}_D[x^2 + y^2] + \mathbb{E}_D([w_R - x]^2 + [z_R - y]^2)\}. \end{aligned}$$

Then,

$$\begin{aligned} \frac{dG(w)}{dw} = {} & -\frac{2(1+\beta_D^2)}{2\alpha} w\{-[1+\beta_D^2]z^2 + 2\mathbb{E}_D y[1+\beta_D^2]\, z - \mathbb{E}_D[x^2+y^2] + \mathbb{E}_D([w_R - s]^2 + [z_R - y]^2)\} + \\ & 2(1+\beta_D^2)[1 - F([1+\beta_D^2]z^2 - [w_R^2 + z_R^2])][\mathbb{E}_D\, x - w]. \end{aligned} \tag{29}$$

Since $(\hat{w}, \hat{z})$ satisfies (16), by substituting $\beta_D \hat{w}$ for $\hat{z}$ in (16) we may verify that $\hat{w}$ satisfies:

$$\frac{dG(w)}{dw}|_{w=\hat{w}} = 0.$$

Next, we show $d^2G(w)/dw^2 < 0$ for all $(w, \beta_D w) \in C_D$. Differentiating (29) yields:

$$\begin{aligned}
&\frac{dG^2}{dw^2} \propto \\
&-\frac{2}{\alpha}\{-[1+\beta_D^2]z^2 + 2\mathbb{E}_D y[1+\beta_D^2]\,z - \mathbb{E}_D[x^2+y^2] + \mathbb{E}_D([w_R - x]^2 + [z_R - y]^2)\} \\
&-\frac{2(1+\beta_D^2)}{\alpha} z\{\mathbb{E}_D y - z\} \\
&-2(1+\beta_D^2)[1 - F([1+\beta_D^2]z^2 - [w_R^2 + z_R^2])]. \qquad (30)
\end{aligned}$$

To sign this expression we sign each of the expressions in the second through the fourth line. In the second line, the expression in $\{\cdot\}$ is the net gain from winning the election. By condition (ii) in the statement of the proposition, $(\beta_D z, z) = (0, 0)$ is at least as good as (w_R, z_R). Since the net gain from winning the election is increasing in z as z varies from 0 toward $\mathbb{E}_D y$, $\{\cdot\} \geq 0$ and, so, $-(2/\alpha)\{\cdot\} \leq 0$. Turning to the third line, observe that as z varies from 0 toward $\mathbb{E}_D y$, $z[\mathbb{E}_D y - z] \geq 0$ regardless of the sign of $\mathbb{E}_D y$ and, so, the expression in the third line is less than or equal to 0. In the final line, the term $[1 - F(\cdot)] \in (0, 1)$ and so the expression is strictly less than zero. Hence $d^2G^/dz^2 < 0$ for all z such that $(\beta_D z, z) \in C_D$. It follows that $(\hat{w}_k, \hat{z}_k)$ strictly dominates any other point in C_D.

To complete the proof, we need to show that $(\hat{w}_D, \hat{z}_D)$ also dominates any point outside of C_D. Let $T = \{(w, z) : w^2 + z^2 - [w_R^2 + z_R^2] \geq \alpha\}$. For $(w, z) \in T$, the D party loses the election for sure and, so, the payoff from such points is $-L(w_R, z_R)$. Since $(w, z) = (w_R, z_R)$ belongs in C_D, and the payoff to party D from this choice is $-L(w_R, z_R)$, it follows that the payoff from $(\hat{w}_D, \hat{z}_D)$ strictly dominates the payoff from any choice in T.

Next, let $B = \{(w, z) : w^2 + z^2 - [w_R^2 + z_R^2] \leq -\alpha\}$. If $w_R^2 + z_R^2 < \alpha$, B is empty and there is nothing to check. Assume, then, that $w_R^2 + z_R^2 \geq \alpha$. For $(w, z) \in B$, the D party wins the election for sure. In this set, the best that party D can do is to set $w = \beta_D z$ and pick z to be as close as possible to $\mathbb{E}_D y$. This choice will put the party on the border separating B from C_D. By continuity of the objective function, a point on the border will give essentially the same payoff as a point that is arbitrarily close to it but in C_D. Since $(\hat{w}_D, \hat{z}_D)$ strictly dominates any other point in C_D, it must strictly dominate any point on the border as well.

Hence $(\hat{w}_D, \hat{z}_D) = (w_D^*, z_D^*)$.

B Convergent Equilibrium

Thus far we have focused on polarized equilibria and these will be the main focus of the paper. For completeness, we describe the case of convergent equilibria.

To start off, observe that (3) implies that if both parties have the same platform, then a person's vote is entirely determined by the realization of A. All voters vote for the D party if $A \geq 0$; otherwise all vote for the R party. Thus, in the convergent case, $\bar{A}(\mathcal{P}) = 0$ and $\mathbb{A}(\mathcal{P}) = \{A \in \mathbb{R} : A \geq 0\}$. Furthermore, if the two parties are indistinguishable for voters then every (x, y) belongs to both $H_D(\mathcal{P}^e)$ and $H_R(\mathcal{P}^e)$. Hence, $H_k(\mathcal{P}^e) = \mathbb{R}^2,\ k \in \{D, R\}$.

Focusing on symmetric convergent equilibria, the only equilibrium configuration possible is $(0, 0, 0, 0)$. The next two propositions together prove this.

Proposition B.0.1. *Let* (w^*, z^*) *be party* k*'s best response to* $(w_{\sim k}, z_{\sim k})$*. If* $\mathbb{E}(x|H_k(\mathcal{P}^e)) = 0$ *and* $\mathbb{E}(y|H_k(\mathcal{P}^e) = 0$ *then* $(w^*, z^*) = (0, 0)$*.*

Proof. We will prove this for the D party. For any $(w, z) \neq (0, 0)$, from (14) we have $\mathbb{E}_{(x,y)|H_k(\mathcal{P}^e)}\{\cdot\}$ is strictly smaller and the probability of winning not any greater. Thus, for any $(w, z) \neq (0, 0)$ the objective function is strictly smaller than if $(w, z) = (0, 0)$. Hence, $(w^*, z^*) = (0, 0)$ for any $(w_{\sim k}, z_{\sim k})$. □

Proposition B.0.2. $\mathcal{P} = (0, 0, 0, 0)$ *is an equilibrium.*

Proof. Let $\mathcal{P}^e = (0, 0, 0, 0)$. Since beliefs are convergent, $H_k(\mathcal{P}^e) = \mathbb{R}^2$ for $k \in \{D, R\}$. By the symmetry of the (x, y) distribution, $\mathbb{E}(x|H_k(\mathcal{P}^e)) = \mathbb{E}(y|H_k(\mathcal{P}^e)) = 0$, $k \in \{D, R\}$. By Proposition B.0.1, $(w_k^*, z_k^*) = (0, 0)$ for $k \in \{D, R\}$. Since outcomes given beliefs coincide with beliefs, $\mathcal{P} = (0, 0, 0, 0)$ is an equilibrium. □

Thus the Black-Downs-Hotelling "principle of minimum differentiation" holds in our environment, i.e., there is an equilibrium in which parties choose the same policies and the policies are the mean of the voters' ideal policies.