Inequality, Participation, and Polarization:

Economic Origins of Partisan Policies

Razvan Vlaicu

Department of Research and Chief Economist

DISCUSSION PAPER Nº IDB-DP-1040

November 2023

Inequality, Participation, and Polarization:

Economic Origins of Partisan Policies

Razvan Vlaicu

Inter-American Development Bank

November 2023

Inter-American Development Bank

Inequality, Participation, and Polarization: Economic Origins of Partisan Policies*

Razvan Vlaicu†

November 2023

Abstract

The upward co-movement of income inequality and partisan polarization in the U.S. is typically attributed to intensified class conflict or a political wealth bias. This paper formalizes a theory of polarization where changes in the income distribution do not affect citizens' policy preferences, but instead change their patterns of political participation: aggregate voting decreases relative to aggregate giving, reducing the electoral penalty for partisan policies. By endogenizing party composition the model captures both the ideological and compositional dimensions of polarization, and addresses less-discussed polarization features, such as intra-party homogeneity and the increase in safe seats. According to the model, observed polarization patterns imply that parties have diverged more than candidates, and that the gap between party and candidate divergence has increased with income inequality.

JEL Classification: D72, D78, H11.

Keywords: income inequality, electoral competition, political participation, partisan polarization.

*I am grateful to Theodore Bergstrom, Dan Bernhardt, Georgy Egorov, Tim Feddersen, Laurel Harbridge, Ethan Kaplan, Phil Keefer, Nolan McCarty, Daven Petitte, Mattias Polborn, Richard Van Weelden, Emanuel Vespa, and seminar audiences at the IDB Research Department, Northwestern University, UC Santa Barbara, University of Chicago, and the University of Maryland. Gregory Haugan provided excellent research assistance. The findings and interpretations in this paper are those of the author and do not necessarily reflect the views of the Inter-American Development Bank or the governments it represents.

†Research Department, Inter-American Development Bank, 1300 New York Ave NW, Washington, DC 20577, USA. Email: vlaicu@iadb.org.

1 Introduction

The steady increase in partisan polarization in the U.S. in the second half of the twentieth century has been a significant and much discussed development. First documented by Poole and Rosenthal (1984), the trend has continued unabated into the beginning of the new century. The first panel of Figure 1 presents the time series of partisan polarization in the House and Senate post World War II. Partisan polarization is measured as the difference between the average ideological scores of Republicans and Democrats (Poole and Rosenthal 1997).[1] According to this measure, in the 113th Congress (2013-2015) the parties were further apart ideologically than at any time not only since World War II but also since the Civil War. As partisan polarization has been shown to negatively affect economic policy and macroeconomic performance (Azzimonti 2015), it is important to understand the electoral foundations of this phenomenon.[2]

The rise in partisan polarization has been accompanied by a rise in income inequality. From 1947-2015 the Gini coefficient for family income climbed from 0.376 to 0.448, an almost 20% increase; see the second panel of Figure 1. Interestingly, the economic and the political series moved in tight lockstep with each other. In Figure 1 the correlation between partisan polarization in the U.S. House and the Gini coefficient is 0.96. The correlation between the detrended versions of these two series is 0.81.[3]

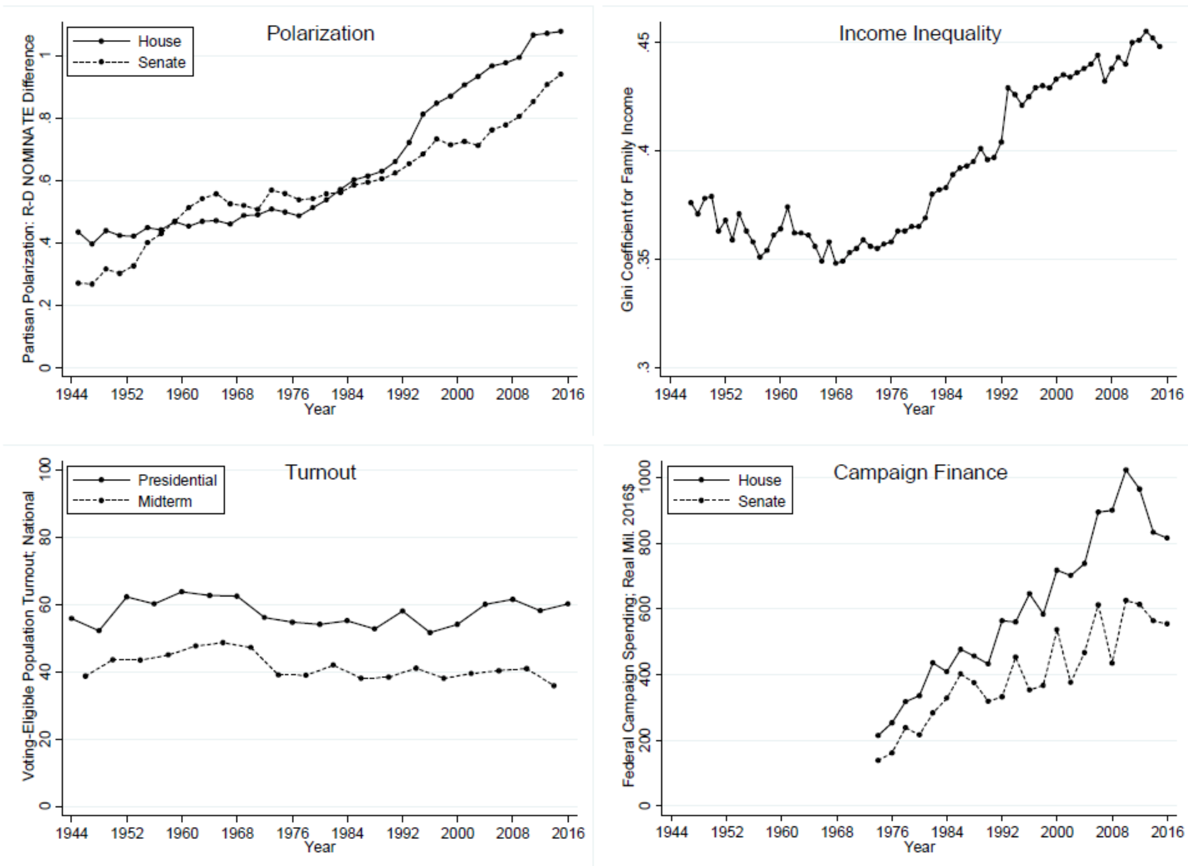
Citizen-level data does not, however, seem to reveal changes in citizen ideology commensurate with the pronounced upward co-movement of inequality and polarization. Ideological polarization in the electorate has been less pronounced and more stable than the ideological polarization observed in Congress (Fiorina and Abrams 2008), although others see a polarization among the politically engaged (Abramowitz 2010). However, to the extent that the politically engaged segment of the public has polarized, or has become more clearly sorted into distinct partisan identities, the timing of the changes seems to indicate that the politicians polarized first and the public afterward (Hetherington 2001, Levendusky 2009).

[1] Ideological DW-NOMINATE scores are derived from scaling methods that use all roll call votes in the U.S. Congress and assume a spatial model of voting.

[2] However, see Van Weelden (2013) for a model where divergent platforms can be welfare enhancing to a representative voter, and Van Weelden (2015) for conditions under which platform divergence can be socially optimal for a heterogenous electorate.

[3] The correlation is also present in cross-section: across senators (Garand 2010) and across state legislatures (Shor and McCarty 2011). Time series correlations using longer, pre-1947, series of income inequality, e.g., the share of income going to the top 1% earners (Piketty and Saez 2003), or the inverted Pareto-Lorenz coefficient (Duca and Saving 2016), yield similar magnitudes.

Figure 1: Time Trends.



Notes: Data sources are as follows. Polarization: Keith Poole's voteview.com. Gini coefficient: U.S. Census Bureau. Turnout: Michael McDonald's electionproject.org. Campaign spending: The Campaign Finance Institute data based on Federal Election Commission (FEC) files. The Census Bureau has been reporting the Gini coefficient since 1947. The FEC has been reporting campaign finance data starting in 1974. Campaign spending roughly equals aggregate individual contributions since both are about 60% of total campaign funds raised.

Interestingly, voter turnout in federal elections has not increased despite sharper differences between candidates and parties. Importantly, low turnout has coincided with a surge in campaign spending, one of whose main goals is arguably to mobilize voters (Herrera, Levine, and Martinelli 2008, p. 503); see bottom panels of Figure 1. Thus, the voluntary, counterfactual, component of turnout, i.e., turnout in the absence of mobilization efforts, is likely to have declined. By contrast, the number of individuals making itemized contributions has grown by a factor of about ten since 1980; contributions from both large and small donors go fairly evenly to Democrats and Republicans (Bonica, McCarty, Poole, and Rosenthal 2013). The salience of income redistribution in political discourse has diminished (Gerring 1998), and the salience of cultural issues has increased at the expense of economic issues (Krasa and Polborn 2014a). In the cross-section, a strong association between state income inequality and senator polarization remains even after controlling for citizen ideology and citizen polarization (Garand 2010). Overall, the inequality-polarization link appears to have empirical features that cannot be fully explained by increased citizen polarization over economic issues.

In this paper I formalize a theory of polarization where changing income inequality does not affect citizens' policy preferences. Instead, it affects their political participation. I note that while income as a predictor of ideology has been subject to much debate (see, e.g., Jacobson 2012), income as a predictor of participation has proven more robust, e.g., voting, giving, and the share of income given, have been shown to strongly correlate with income. Consequently I turn my attention to the effect of income inequality on participation. To study the participation channel in isolation, I develop a model where citizens' ideology is orthogonal to their income. However, citizens' propensity to vote and to give to parties and candidates depends on their income. Parties and candidates are policy-motivated. Parties choose ranges of acceptable policies, and candidates choose unique policy positions. Both choices are made under aggregate uncertainty about the mean ideology of political donors.[4]

The basic mechanism is as follows. Because voting propensity is more sensitive to an income change for low-income citizens, and giving propensity more sensitive to an income change for high-income citizens, income inequality, by reducing incomes at the bottom and raising them at the top, depresses aggregate voting and bolsters aggregate giving. This makes election results relatively more dependent on donor support. Because mean donor policy preferences are uncertain, candidates' winning probabilities are now more loosely linked to their policy positions. This lowers candidates' electoral cost of pursuing partisan policy goals at the expense of centrist citizens in their district. Parties choose policy intervals that determine which candidates affiliate with them, thus party composition is endogenous. As parties also face donor uncertainty, parties have stronger incentives to diverge the higher is aggregate giving. Income inequality, by elevating the importance of giving, thus leads to both higher candidate divergence and higher party divergence.

Candidate divergence and party divergence are the two mechanisms that create partisan polarization in the model. *Partisan polarization* is defined as the difference between the average policy position of right party candidates and the average policy position of left party

[4] In this paper I focus on individual contributions. These comprise the majority of all political donations in U.S. federal elections. An interesting extension would be to introduce interest-group contributions.

candidates. The two mechanisms allow the decomposition of partisan polarization into its ideological and compositional components. *Ideological polarization* is a shift in candidate policy positions without a change in candidate partisan affiliations. *Compositional polarization* is a change in candidate partisan affiliations without a shift in candidate policy positions. I show that the former equals candidate divergence, while the latter is proportional to the gap between party and candidate divergence.

The model predicts the observed correlation between income inequality and partisan polarization, but also matches polarization data patterns that have received less attention though are equally significant features of the U.S. polarization experience. Specifically, it predicts that income inequality, by increasing the importance of giving relative to voting, increases inter-party heterogeneity and the prevalence of safe seats, and decreases intra-party heterogeneity and party overlap. The model demonstrates that these features reflect a polarization driven by parties rather than candidates. In other words, it implies that parties have diverged more than candidates, and that the gap between party and candidate divergence has increased with income inequality. The model also nests exogenous factors that have been previously linked to polarization, such as within- and across-district ideological heterogeneity, and shows that, unlike income inequality, these variables cannot provide a full account of the polarization features observed in the data.

Inequality-polarization correlations and a finding that partisan stratification by income has increased after the 1970s have led McCarty, Poole, and Rosenthal (2006) to conjecture that electoral politics in the U.S. have become more class-based. That is, as the rich become richer and the poor poorer, shrinking the size of the middle class, the citizenry becomes more polarized into economic classes, with more clearly differentiated policy preferences. Low-income voters are increasingly represented by liberal (pro-redistribution) candidates while high-income citizens are increasingly represented by conservative (anti-redistribution) candidates. An alternative explanation of the inequality-polarization link is a political "wealth bias." Bartels (2008) provides evidence that while political parties have polarized they have at the same time both moved to the right of the median voter. He argues that a key factor has been the increasing influence of conservative political donations. Feddersen and Gul (2015) formalize this view by assuming that candidates pursue both voters and donors; donors are biased against redistribution and more so when inequality increases. Candidates are uncertain about the relative importance of votes vs. money so in equilibrium candidates polarize, with each one closer to one of the two constituencies.[5] By contrast, my model removes the possibility of a class conflict or of a wealth bias by assuming that (i) voter and donor policy preferences are invariant to income inequality, and (ii) political competitors have equal donor support ex ante. This allows me to isolate the participation channel. Thus, my model complements the existing explanations and accounts for the apparent "disconnect" between partisan polarization and citizen preferences.[6]

Several theoretical approaches have been proposed to model electoral divergence: policy-motivated candidates with uncertainty (Wittman 1983, Calvert 1985), entry-deterrence of third parties in a single district (Palfrey 1984) or heterogeneous districts (Callander 2005), candidates with fixed and differentiated characteristics (Krasa and Polborn 2010), citizen-candidates with rent-seeking opportunities in office (Van Weelden 2013), probabilistic voting by a polarized electorate with convex preferences (Kamada and Kojima 2014), quadratic voting (Patty and Penn 2017). These models generally assume no abstention and no individual contributions. I build on the Calvert-Wittman framework by allowing aggregate participation - voting and giving - to vary with the income distribution. Another innovation of the model is the link between parties and candidates. Parties compete across heterogenous districts for candidates who compete in their own districts. Introducing parties with endogenous composition is key because it allows the identification of the ideological and compositional dimensions of partisan polarization, and captures polarization features beyond mean ideological differences.[7]

The paper is organized as follows. Section 2 introduces the model. Section 3 characterizes its equilibrium and derives comparative statics with respect to inequality. Section 4 draws implications about the mechanisms driving U.S. polarization and considers alternative explanations. Section 5 concludes and discusses possible extensions. Proofs of all the formal results are contained in the Appendix.

---

[5] Using a more reduced-form political process, Campante (2011) shows that income inequality may increase the influence of individual donors on the election outcome, reducing the political power of the median voter and resulting in less redistributive positions.

[6] Großer and Palfrey (2014) provide a different rationale for the "disconnect" between partisan polarization and citizen polarization, based on candidate risk-aversion, using a citizen-candidate model with incomplete information.

[7] To my knowledge, this is the first paper that endogenizes internal party composition in a Calvert-Wittman framework. In this respect the paper is related to a small literature that has explicitly modeled the link between parties and candidates, e.g., Austen-Smith (1984), Snyder and Ting (2002). In both papers parties are purely office-motivated, rather than policy-motivated as here. Platform divergence is driven by either electoral or signaling considerations.

2 Model

The model is inspired by the class of one-dimensional one-district models of political competition with two policy-motivated candidates competing under aggregate electoral uncertainty (Wittman 1983, Calvert 1985). The novel features of the current setting are (i) income-based citizen participation through voting and giving, and (ii) political parties competing for candidates across multiple, ideologically overlapping, districts.

Consider a continuum of districts, each of which contains a continuum of citizens who differ in ideology z and income y. The district mean ideology will be used to index districts and is distributed uniformly $s \sim U\left[-\frac{\mu}{2}, \frac{\mu}{2}\right]$. Thus, μ is a measure of across-districts ideological heterogeneity. In a given district s, citizen characteristics are distributed according to a joint cdf: $(z, y) \sim F_s\left(\mathbb{R} \times \mathbb{R}_+\right)$. Income is orthogonal to ideology at the district level: $F_s(z, y) = Z_s(z)\, Y_\lambda(y)$, where λ measures income inequality. All districts have the same income distribution and mean income is normalized to unity: $\int_0^\infty y dY_\lambda(y) = 1$, for all λ. If Y_λ is continuously differentiable, then income inequality can be defined as the Gini coefficient using: $\lambda \equiv \int_0^\infty Y_\lambda(y)\left[1 - Y_\lambda(y)\right] dy$. Also, an increase in income inequality can be captured as a mean-preserving spread of the current income distribution.[8]

Two candidates compete in each district, and two parties compete for seats across districts. The timing of the game is the following:

(I) The parties $\mathcal{L}, \mathcal{R}$ know the distribution of district ideological means, are uncertain about the mean of party donor support, and simultaneously commit to policy intervals $X_{\mathcal{L}} \pm \frac{\mu}{2}, X_{\mathcal{R}} \pm \frac{\mu}{2}$ around their party positions $X_{\mathcal{L}}, X_{\mathcal{R}}$.

(II) In each district $s \in \left[-\frac{\mu}{2}, \frac{\mu}{2}\right]$ the opposing candidates L_s, R_s know the distribution of electoral support, are uncertain about the mean of donor support, and simultaneously commit to policy positions x_{L_s}, x_{R_s}, which are points on the real line.[9]

Parties are policy motivated, in the sense that they care about the policy adopted by the legislature formed as a result of the elections. Party $\mathcal{L}$ prefers a smaller legislative policy to a larger one, party $\mathcal{R}$ prefers a larger legislative policy to a smaller one. The legislative policy depends on the parties' policy positions in the election $X_{\mathcal{L}}, X_{\mathcal{R}}$, as well as the parties' legislative strength, or bargaining power in the legislature, denoted by $W_{\mathcal{L}}, W_{\mathcal{R}}$, which in

[8] Formally, $\lambda_1 < \lambda_2$ if and only if $Y_{\lambda_1}(y)$ second-order stochastically dominates $Y_{\lambda_2}(y)$, i.e., $\int_{y<\hat{y}} Y_{\lambda_1}(y)dy \leq \int_{y<\hat{y}} Y_{\lambda_2}(y)dy$, for all $\hat{y}$ in the income support.

[9] In the context of U.S. politics, the candidate labels L, R could be interpreted as liberal and conservative, and the party labels $\mathcal{L}, \mathcal{R}$ as Democrat and Republican.

turn depend on the parties' seat shares and their donations. A party's objective function is:

$$U_p = \begin{cases} -X_{\mathcal{L}}W_{\mathcal{L}} - X_{\mathcal{R}}W_{\mathcal{R}} & \text{if } p = \mathcal{L} \\ X_{\mathcal{L}}W_{\mathcal{L}} + X_{\mathcal{R}}W_{\mathcal{R}} & \text{if } p = \mathcal{R} \end{cases} \tag{1}$$

The legislative outcome is thus a weighted average of the two parties' policy positions. Candidates are also policy motivated, in the sense that they care about the policy position of the winning candidate in their district. A candidate's objective function is:

$$u_{k_s} = \begin{cases} -x_{L_s}w_{L_s} - x_{R_s}w_{R_s} & \text{if } k_s = L_s \\ x_{L_s}w_{L_s} + x_{R_s}w_{R_s} & \text{if } k_s = R_s \end{cases} \tag{2}$$

Here w_{k_s} denotes candidate k_s's winning probability. Thus, an L candidate is better off the more leftward is the expected policy, and an R candidate is better off the more rightward is the expected policy.[10]

Citizen political preferences are captured in reduced form by ideological and donor support for their district candidates, and their donor support for the parties. The cdf of *ideological support* is logistic with mean s and standard deviation scale parameter $\sigma_v > 0$. Denote it by $Z_s(z) = \left[1 + e^{-\frac{z-s}{\sigma_v}}\right]^{-1}$, where $z \in \mathbb{R}$. Then σ_v is a measure of within-district ideological heterogeneity. The cdf of *donor support* is also logistic, with mean $s + \tilde{\psi}_s$ and standard deviation scale parameter $\sigma_g > 0$. Denote it by $D_s\left(z|\tilde{\psi}_s\right) = \left[1 + e^{-\frac{z-s-\tilde{\psi}_s}{\sigma_g}}\right]^{-1}$, where $\tilde{\psi}_s \overset{\text{iid}}{\sim} U\left[-\frac{\psi}{2}, \frac{\psi}{2}\right]$. The parameter ψ measures the candidates' uncertainty about the mean ideology of the candidates' donors. Donor support for parties is $D\left(z|\tilde{\Psi}\right) = \left[1 + e^{-\frac{z-\tilde{\Psi}}{\Sigma_g}}\right]^{-1}$, where $\tilde{\Psi} \sim U\left[-\frac{\Psi}{2}, \frac{\Psi}{2}\right]$. The parameter Ψ measures the parties' uncertainty about the mean ideology of the parties' donors. Under U.S. campaign finance laws, a citizen can donate both to candidates and to parties. Hence the assumption that the distributions of donor support for candidates respectively to parties, D_s, D, are distinct objects. Total contributions to parties have roughly equaled total contributions to candidates, estimated to have been around \$700m in the 1999-2000 election cycle (Ansolabehere, de Figueiredo, and Snyder 2003).[11]

Note that mean ideological support is assumed known to be s, whereas mean donor

[10] Equivalently, one could assume that the candidates L_s, R_s have ideal points located at $s - \theta$ and $s + \theta$, respectively, for some positive θ.

[11] As will be shown below, the logistic functional form helps sidestep thorny equilibrium existence issues common in the Calvert-Wittman framework; see also Feddersen and Gul (2015).

support is uncertain, equal to $s + \tilde{\psi}_s$. In other words, candidates know the mean voter ideology in each district, but are uncertain about the mean donor ideology (or more precisely, mean dollar ideology, if donors vary in the amounts given) in the district. This seems reasonable as predicting electoral support requires predicting citizens' preferences, while predicting donor support requires predicting both donors' preferences as well as the size of their donations. In General Social Survey (GSS) data, between 1980-2014 mean voter ideology on a $[-1, 1]$ scale had a range of variation of 0.076. During the same period, in the Database on Ideology, Money in Politics, and Elections (DIME) based on FEC disclosures, mean (donations-weighted) donor ideology, normalized to the same $[-1, 1]$ scale, had a range of variation of 0.121 (House) and 0.178 (Senate).[12] Also, since 1980, the annual growth rate in the number of voters was around 1.2%, whereas the annual growth rate in individual donors was about 6.6%. These empirical patterns suggest that it is more difficult to predict the ideology of the mean dollar than of the mean voter firstly, because the former varies more from year to year, and secondly, because the population of donors changes more quickly than the population of voters.[13]

Citizens are expressive, and their political participation through voting and giving depends on their income. Let $\nu(y)$ denote a citizen's voting probability, i.e., the likelihood the citizen turns out to vote based on personal income, unaffected by candidates' mobilization efforts. Let $\gamma(y)$ denote a citizen's giving probability. Assume that $\nu(y)$ is strictly concave in income and $\gamma(y)$ is strictly convex in income. Thus, voting is more responsive to an income decrease for lower-income citizens while giving is more responsive to an income increase for higher-income citizens. The first assumption is consistent with findings by Ansolabehere and Hersh (2012) that while self-reported voting is approximately linearly increasing in income, over-reporting is sizable and also increasing in income. The second assumption captures the notion that political giving is a normal good, and individuals' willingness to contribute increases in their income (Campante 2011).[14]

[12] Voter ideology is self-reported on a discrete 1-7 liberal-conservative scale. Donor ideology is estimated by Bonica (2014) on a continuous $[-2, 2]$ interval. The statistics presented use both variables rescaled to the common interval $[-1, 1]$.

[13] One might argue that some large donors' preferences and outlays are easier to predict than many voters' preferences. However, large donors give in multiple races, thus in a given district which candidate will attract more donations may still be difficult to predict since it depends on what happens in all other races.

[14] There is also a well-documented positive association between income and individual voting. Introducing this monotonicity here is not necessary as what matters in the model is aggregate, rather than individual, voting, and aggregate (average) income is fixed to focus on inequality. Individual voting may also increase in ideological extremism, as citizens with stronger preferences are more politically engaged (Jacobson 2012).

Aggregate voting and giving in each district are obtained by integrating across income levels:

$$v(\lambda) = \int_0^\infty \nu(y)\,dY_\lambda(y) \quad \text{and} \quad g(\lambda) = \int_0^\infty \gamma(y)\,dY_\lambda(y) \tag{3}$$

For clarity of interpretation, it is important to emphasize that aggregate voting, as defined in equation (3), should be regarded as counterfactual, or ex ante, turnout, i.e., citizens' voluntary propensity to vote before candidates start deploying their resources to boost turnout. Note that individual participation depends on income y, an individual characteristic, while aggregate participation depends on income inequality λ, a feature of the income distribution.

Candidates' winning probabilities are given by:

$$w_{L_s}(\overline{x}_s, \lambda) = \mathbb{P}\left\{ \frac{\left[Z_s(\overline{x}_s)\right]^{\frac{v(\lambda)}{v(\lambda)+g(\lambda)}} \left[D_s\left(\overline{x}_s|\tilde{\psi}_s\right)\right]^{\frac{g(\lambda)}{v(\lambda)+g(\lambda)}}}{\left[1 - Z_s(\overline{x}_s)\right]^{\frac{v(\lambda)}{v(\lambda)+g(\lambda)}} \left[1 - D_s\left(\overline{x}_s|\tilde{\psi}_s\right)\right]^{\frac{g(\lambda)}{v(\lambda)+g(\lambda)}}} \geq 1 \right\} \tag{4}$$

and $w_{R_s}(\overline{x}_s, \lambda) = 1 - w_{L_s}(\overline{x}_s, \lambda)$, with $\overline{x}_s \equiv \frac{x_{L_s}+x_{R_s}}{2}$. The fraction inside the probability is a standard contest function, whose numerator is the turnout in favor of candidate L_s, and the denominator is the turnout in favor of candidate R_s. A candidate's turnout depends on their ideological support in the electorate, as all potential voters with $z < \overline{x}_s$ would support candidate L_s, and the candidate's (uncertain) donor support, as all potential donors with $z < \overline{x}_s$ would support candidate L_s. Each of these factors' impacts depend on aggregate participation in the form of voting $v(\lambda)$ and giving $g(\lambda)$. The greater aggregate voting $v(\lambda)$, all else equal, the more winning probabilities depend on ideological support $Z_s(\overline{x}_s)$. The greater aggregate giving $g(\lambda)$, all else equal, the more winnong probabilities depend on donor support $D_s\left(\overline{x}_s|\tilde{\psi}_s\right)$.[15]

Which party a candidate belongs to depends on the candidate's policy position x_{k_s} relative to the parties' positions $X_{\mathcal{L}}, X_{\mathcal{R}}$. A candidate k_s affliliates with the party in whose policy interval that candidate's policy position x_{k_s} lies. If a candidate's policy position lies in both parties' policy intervals, the candidate is affiliated with the party whose ideological bent the

Embedding this empirical feature in the model would further increase candidates' incentives for divergence.

[15] Using data from the 2004 and 2008 elections, Spenkuch and Toniatti (2016) find that differences in advertising between the two parties across different media markets lead to differences in the parties' vote shares. Advertising accounts for about half of campaign spending; other campaign activities such as canvassing and direct mail have also been shown to affect turnout (Gerber and Green 2000).

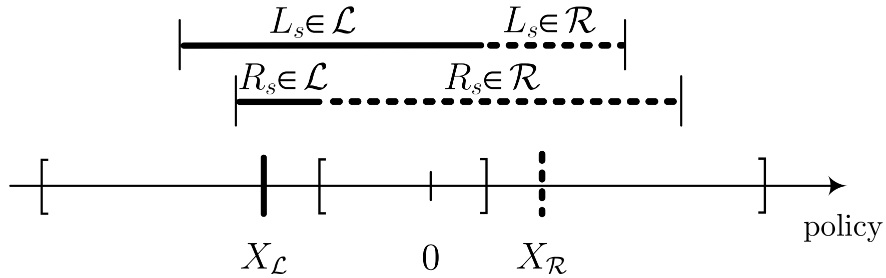


Figure 2: Candidates' Party Affiliations; Non-Symmetric Scenario.

candidate shares. These assumptions can be formalized as follows.

$$k_s \in \begin{cases} \mathcal{L} & \text{if } x_{k_s} \in \left[X_{\mathcal{L}} \pm \frac{\mu}{2}\right] \setminus \left[X_{\mathcal{R}} \pm \frac{\mu}{2}\right] \\ \mathcal{L}\mathbf{1}_{\{k_s=L_s\}} + \mathcal{R}\mathbf{1}_{\{k_s=R_s\}} & \text{if } x_{k_s} \in \left[X_{\mathcal{L}} \pm \frac{\mu}{2}\right] \cap \left[X_{\mathcal{R}} \pm \frac{\mu}{2}\right] \\ \mathcal{R} & \text{if } x_{k_s} \in \left[X_{\mathcal{R}} \pm \frac{\mu}{2}\right] \setminus \left[X_{\mathcal{L}} \pm \frac{\mu}{2}\right] \end{cases} \tag{5}$$

Figure 2 illustrates the affiliation rule. Note that party $\mathcal{L}$ candidates (bold solid lines) in some districts can be rightward of party $\mathcal{R}$ candidates (bold dashed lines) in other districts, e.g., Texas Democrats more conservative than Massachusetts Republicans. Also, notice that extreme districts may be dominated by one party, as both candidates in these districts are affiliated with the same party; below, these districts are referred to as safe seats.[16]

A party's seat share is the fraction of districts won by its candidates, denoted by $h_p \equiv \frac{1}{\mu}\int_{k_s \in p} \mathbf{1}_{win(k_s)} ds$, for $p = \mathcal{L}, \mathcal{R}$. Parties' policy weights in the legislature depend on relative seat shares and donations to parties as follows:

$$W_{\mathcal{L}} = \mathbb{P}\left\{ \frac{\left[Z\left(h_{\mathcal{L}} - h_{\mathcal{R}}\right)\right]^{\frac{1}{1+g(\lambda)}} \left[D\left(\bar{X}|\tilde{\Psi}\right)\right]^{\frac{g(\lambda)}{1+g(\lambda)}}}{\left[1 - Z\left(h_{\mathcal{L}} - h_{\mathcal{R}}\right)\right]^{\frac{1}{1+g(\lambda)}} \left[1 - D\left(\bar{X}|\tilde{\Psi}\right)\right]^{\frac{g(\lambda)}{1+g(\lambda)}}} \geq 1 \right\} \tag{6}$$

and $W_{\mathcal{R}} = 1 - W_{\mathcal{L}}$, with $\bar{X} \equiv \frac{X_{\mathcal{L}} + X_{\mathcal{R}}}{2}$. Legislative support for party $\mathcal{L}$'s policy position is assumed to depend on its seat share margin $h_{\mathcal{L}} - h_{\mathcal{R}}$, according to the functional form $Z\left(h_{\mathcal{L}} - h_{\mathcal{R}}\right) = \left[1 + e^{-\frac{h_{\mathcal{L}} - h_{\mathcal{R}}}{2}}\right]^{-1}$. The numerator of the fraction inside the probability is

[16]The tie-breaking assumption rules out situations where an L candidate affiliated with party $\mathcal{R}$ runs against an R candidate affiliated with party $\mathcal{L}$, and can be relaxed without affecting the main results. The affiliation rule assumed here is agnostic about candidates' motivation in adopting a party label. Its role is simply to create a correlation between a party and its candidates' policy positions.

the legislative strength of party $\mathcal{L}$, and the denominator is the legislative strength of party $\mathcal{R}$. These in turn are functions of each party's legislative support and (uncertain) donor support. A party can thus compensate for a seat disadvantage by building a better internal organization using its donations. The relative impacts of legislative support and donor support to parties depend on the level of aggregate giving $g(\lambda)$. The higher aggregate giving $g(\lambda)$ the more party legislative strength depends on party donor support $D\left(\bar{X}|\tilde{\Psi}\right)$, and less on legislative support $Z(h_{\mathcal{L}} - h_{\mathcal{R}})$.[17]

Throughout the rest of the paper the model's parameters are assumed to satisfy:

$$\frac{\psi}{1+\frac{v(\lambda)/\sigma_v}{g(\lambda)/\sigma_g}} \leq \frac{\Psi}{1+\frac{1/\mu}{g(\lambda)/\Sigma_g}} \leq \mu \tag{7}$$

These conditions rule out candidates who run unaffiliated and ensure there is equilibrium ideological overlap between the set of L candidates and the set of R candidates, as well as between the two parties' policy intervals, for all levels of inequality λ.

The equilibrium concept is subgame-perfect Nash equilibrium. Since citizens do not behave strategically, the equilibrium can be stated in terms of party and candidate strategies. An equilibrium is a collection of strategies $\left\{(X_{\mathcal{L}}, X_{\mathcal{R}}), (x_{L_s}, x_{R_s})_{s\in\left[-\frac{\mu}{2},\frac{\mu}{2}\right]}\right\}$, where the two parties $\mathcal{L}, \mathcal{R}$ choose mutual best response policy positions, and each pair of opposing candidates L_s, R_s choose mutual best response policy positions, given parties' strategies.

3 Equilibrium Analysis

In this section I characterize the equilibrium of the model and study its properties. The first result characterizes the link between income inequality and aggregate participation. I then analyze the incentives of candidates competing in district elections, followed by the incentives of parties. The ultimate goal of this analysis is to understand the determinants of partisan polarization, defined as follows.

Definition (Partisan Polarization): *Partisan polarization is the difference between the*

[17] Donations to parties are used for party-building activities that support a party's legislative agenda, such as organizing party conventions and promoting party discipline and collective action. Some donations to parties may be transferred to candidates, but for simplicity the model abstracts from this possibility. For models that endogenize the majority party's influence on policymaking in a legislative body see, e.g., Patty (2008), Diermeier and Vlaicu (2011), Diermeier, Prato, and Vlaicu (2015).

average policy position of a party $\mathcal{R}$ member and the average policy position of a party $\mathcal{L}$ member,

$$\Delta\mathbb{E}\left(x\right) \equiv \frac{1}{\mu}\left(\int_{k_s\in\mathcal{R}} x_{k_s}ds - \int_{k_s\in\mathcal{L}} x_{k_s}ds\right) \tag{8}$$

where $\frac{1}{\mu}$ is the density of the distribution of mean ideology s across districts.

Aggregate citizen participation through voting and giving affects both candidates' and parties' payoffs. This is because aggregate voting and giving $v\left(\lambda\right), g\left(\lambda\right)$ translate ideological and donor support into electoral support for candidates and legislative strength for parties. When individual participation depends on income in an environment with constant average income, aggregate participation varies with income inequality as described by the following result.

Proposition 1 (Aggregate Participation): *Assume that individual voting is strictly concave in income, $\frac{d^2\nu(y)}{dy^2} < 0$, and individual giving strictly convex in income, $\frac{d^2\gamma(y)}{dy^2} > 0$. Then aggregate voting is strictly decreasing in income inequality, $\frac{dv(\lambda)}{d\lambda} < 0$, and aggregate giving is strictly increasing in income inequality, $\frac{dg(\lambda)}{d\lambda} > 0$.*

If individual voting $\nu\left(y\right)$ is strictly concave in income then an increase in income inequality, by shifting individuals toward the top and bottom ends of the income distribution, should depress aggregate voting $v(\lambda)$ because individual voting is lower on average in the tails of the income distribution than in the center. If individual giving $\gamma(y)$ is strictly convex in income then an increase in income inequality, by shifting individuals toward the top and bottom ends of the income distribution, should bolster aggregate giving $g(\lambda)$ because individual giving is higher on average in the tails of the income distribution than in the center.[18]

Next consider electoral competition between district s's candidates, L_s and R_s, for a seat. If candidate R_s pursues its preferred policy position and moves away from candidate L_s the

[18] Had one also assumed voting to be strictly increasing in income, then one could claim that aggregate voting at the bottom of the income distribution drops by more than it increases at the top. An immediate corollary of Proposition 1 is relevant to the "puzzle of political participation" first stated by Brody (1978): if income increases individual voting why hasn't aggregate voting in U.S. elections increased in the post-1960 period when real per-capita income had an upward trend? While higher individual income may boost individual voting, and higher average income aggregate voting, the attendant higher income inequality can act as a countervailing force to keep aggregate voting depressed.

marginal change in candidate R_s's payoff is:

$$\frac{\partial}{\partial x_{R_s}} u_{R_s} = w_{R_s} + (x_{R_s} - x_{L_s}) \frac{\partial}{\partial x_{R_s}} w_{R_s} \tag{9}$$

which shows that by pursuing policy goals this candidate trades off winning probability, which declines at rate $x_{R_s} - x_{L_s} \geq 0$, for policy benefits, which increase at rate $w_{R_s} \geq 0$. In this setting, reduced winning odds come from lower expected electoral and donor support.

Using the assumed forms of the citizen and donor support distributions, one can compute candidate R_s's winning probability:

$$w_{R_s}(\overline{x}_s, \lambda) = \frac{1}{2} - \frac{1}{\psi}\left[1 + \frac{v(\lambda)/\sigma_v}{g(\lambda)/\sigma_g}\right](\overline{x}_s - s) \tag{10}$$

and $w_{L_s}(\overline{x}_s, \lambda) = 1 - w_{R_s}(\overline{x}_s, \lambda)$. This equation reflects that candidate R_s has an electoral advantage ($w_{L_s} > \frac{1}{2}$) over candidate L_s whenever they are closer to the district mean ideology s than their opponent, i.e., $\overline{x}_s < s$, or equivalently $|x_{R_s} - s| < |x_{L_s} - s|$. However, candidates care not only about electoral advantage, but also about policy. A candidate's marginal electoral penalty for moving away from the district mean, namely $\frac{\partial}{\partial x_{R_s}} w_{R_s} = -\frac{1}{2\psi}\left[1 + \frac{v(\lambda)/\sigma_v}{g(\lambda)/\sigma_g}\right]$, is lower the lower the ratio $v(\lambda)/g(\lambda)$, i.e., the more important is giving relative to voting. Conversely, a candidate's policy-motivated deviation from the district mean is more profitable the lower the ratio $v(\lambda)/g(\lambda)$. Higher giving (relative to voting), or higher uncertainty ψ about the mean donor ideology, gives candidates more flexibility to pursue their partisan policy preferences by reducing the electoral penalty for doing so.[19]

In a candidate equilibrium both candidates equate the marginal policy gain of a deviation from the center with the marginal electoral penalty, cf. equation (9).

Proposition 2 (Candidate Competition) *There exists a unique candidate equilibrium in each district s. In this equilibrium candidates L_s, R_s adopt distinct policy positions that are symmetric around the district mean, $(x_{L_s}, x_{R_s}) = \left(s - \frac{\Delta x}{2}, s + \frac{\Delta x}{2}\right)$, for all s, where $\Delta x = \frac{\psi}{1+\frac{v(\lambda)/\sigma_v}{g(\lambda)/\sigma_g}}$. Each candidate wins with probability a half. Equilibrium candidate divergence Δx ranges from zero, when voting dominates giving $\frac{v(\lambda)}{g(\lambda)} \to \infty$, to ψ, when giving dominates voting $\frac{v(\lambda)}{g(\lambda)} \to 0$.*

[19] If candidates were just office-motivated, i.e., they didn't care about policy outcomes, only about winning the seat, in this setting neither would find it profitable to move away from the center since the electoral penalty would not be offset by a policy benefit.

To establish that the strategies described in Proposition 2 constitute an equilibrium it is first necessary to rule out local deviations, namely $\frac{\partial}{\partial x_{L_s}} u_{L_s} = \frac{\partial}{\partial x_{R_s}} u_{R_s} = 0$. This yields a unique symmetric strategy pair. Second, global deviations are ruled out by the strict concavity of candidates' payoff functions in their own strategy. Strict concavity results from the linearity of the policy payoffs in candidate strategies, see equation (2), and the linearity of the winning probability function in candidate strategies, see equation (10); the latter follows from the assumed logistic and uniform distributions together with the contest function form of a candidate's turnout. Strict concavity of payoffs in own strategies thus ensures a unique symmetric equilibrium.

From equation (9), and its analog for candidate L_s, and using equation (10), one can solve for equilibrium *candidate divergence*, namely the difference in candidate policy positions:

$$\Delta x \equiv x_{R_s} - x_{L_s} = \frac{\psi}{1 + \frac{v(\lambda)/\sigma_v}{g(\lambda)/\sigma_g}} \tag{11}$$

which is a function of aggregate participation through voting and giving, which in turn depend on income inequality; the degree of uncertainty about the location of the mean donor ideology; and the dispersions of the ideological and donor support distributions. The next result states how these factors affect candidate divergence in district elections.

Proposition 3 (Candidate Comparative Statics) *Under imperfect aggregate participation, $0 < v(\lambda), g(\lambda) < 1$, equilibrium candidate divergence Δx is, all else equal: (i) strictly increasing in income inequality λ, (ii) strictly increasing in the uncertainty ψ about the mean donor ideology, and (iii) strictly increasing in within-district ideological heterogeneity σ_v, and strictly decreasing in the dispersion σ_g of donor support.*

When aggregate voting and giving move in opposite directions, the direction of change in candidate polarization is clear, it moves in the direction of giving. When aggregate voting and giving move in the same direction, the direction of change in candidate polarization is the direction of giving if and only if giving changes more than voting.[20] When income inequality increases, it depresses aggregate voting and bolsters aggregate giving, by Proposition 1, making the electoral outcome more dependent on donor support. But, donor support is a polarizing force in district elections because it makes more salient the uncertainty about the

[20] Thus assuming strict concavity of individual voting in income is unnecessarily strong. All that is needed is that it is less convex than individual giving.

mean donor ideology, which reduces the candidates' marginal electoral penalty from pursuing partisan policy goals. An exogenous increase in the uncertainty ψ has the same effect.[21]

The key message of this result is that candidate divergence varies with income inequality even in an environment where income is not the dimension of political conflict. It does so through the channel of aggregate voting and giving, both of which are affected by the income distribution.[22]

I now turn to the strategies of the parties. One can think of at least two important advantages of explicitly modeling parties as heterogenous collections of candidates. First, candidate divergence need not necessarily translate into partisan polarization, since parties may have incentives to seek out and nominate moderate candidates, or because parties' moderate candidates are consistently more successful in elections against extreme candidates. Second, by endogenizing party composition one can study the determinants of less discussed, but no less interesting, features of the U.S. polarization experience, such as the increase in intra-party ideological homogeneity (McCarty, Poole, and Rosenthal 2006), and the rise in the fraction of safe seats (Abramowitz, Alexander, and Gunning 2006). These types of variables cannot be captured in a standard two-candidate single-district model of electoral competition because they are by nature distributional features.

When parties position themselves in the ideological space in stage I, they anticipate the candidates' electoral strategies in stage II. Because of candidates' symmetric positioning in each district, any given candidate has an equal chance of winning, $w_{k_s}(\overline{x}_s, \lambda) = \frac{1}{2}$ for all districts $s \in \left[-\frac{\mu}{2}, \frac{\mu}{2}\right]$. As electoral uncertainty is independent across districts, i.e., $\tilde{\psi}_s$ are iid, applying the law of large numbers, district-level electoral uncertainty dissipates when aggregated across districts, and parties' vote shares are:

$$h_p \equiv \frac{1}{\mu} \int_{k_s \in p} \mathbf{1}_{win(k_s)} ds = \frac{1}{\mu} \left(\frac{1}{2} \int_{L_s \in p} ds + \frac{1}{2} \int_{R_s \in p} ds \right) \tag{12}$$

for $p = \mathcal{L}, \mathcal{R}$. In words, a party wins half of the districts where only one candidate is affiliated

[21] See Callander and Wilson (2007) for an alternative model where candidate divergence is also negatively correlated with aggregate voting.

[22] There is broad agreement with the notion that the ideological divide between the two major U.S. parties has been widening for several decades. Consensus breaks down, however, over whether there has been a commensurate ideological divide in the American public. Abramowitz (2010): "Polarization in Washington reflects polarization within the public, especially the politically engaged segment of the public." (p. x). Fiorina and Abrams (2008): "It seems reasonable to conclude that the distribution of ideology in the American public has not changed for more than three decades." (p. 571).

with it, and all the districts where both candidates are affiliated with it, $L_s, R_s \in p$. Using the properties of equilibrium candidate competition derived in Proposition 2, the seat share margin can be shown to take the form:

$$h_{\mathcal{L}} - h_{\mathcal{R}} = \frac{1}{\mu}\left(X_{\mathcal{L}} + X_{\mathcal{R}}\right) \tag{13}$$

which shows that on the equilibrium path the symmetry of candidate positions makes the seat share margin independent of candidate divergence. Intuitively, L candidate divergence exactly cancels R candidate divergence.

The model assumes that parties, in a similar way as candidates, care about policy outcomes; see the party objective functions in equation (1). To attain its policy objectives a party needs sufficient legislative strength, which in turn depends on the share of seats it wins, and the amount of donations it receives. Policy goals push parties toward the extremes, while donor support push them toward the center as the distribution of donor preferences is unbiased.

The tradeoff a party has to resolve in equilibrium can be seen in the expression of a party's policy weight in the legislature; refer to equation (6):

$$W_{\mathcal{R}}\left(\bar{X}, \lambda\right) = \frac{1}{2} - \frac{1}{\Psi}\left[1 + \frac{1/\mu}{g\left(\lambda\right)/\Sigma_g}\right]\bar{X} \tag{14}$$

where $\bar{X} \equiv \frac{X_{\mathcal{L}}+X_{\mathcal{R}}}{2}$, and $W_{\mathcal{L}}\left(\bar{X}, \lambda\right) = 1 - W_{\mathcal{R}}\left(\bar{X}, \lambda\right)$. This expression shows that party $\mathcal{R}$ has a legislative advantage ($W_{\mathcal{R}} > \frac{1}{2}$) over party $\mathcal{L}$ whenever it is closer to the mean of district ideological means than its opponent, i.e., $\bar{X} < 0$, or equivalently $|X_{\mathcal{R}}| < |X_{\mathcal{L}}|$. A party's marginal legislative penalty for moving away from center in order to stake a more preferred policy position is $\frac{\partial}{\partial X_{\mathcal{R}}} W_{\mathcal{R}} = -\frac{1}{2\Psi}\left[1 + \frac{1/\mu}{g(\lambda)/\Sigma_g}\right]$. Note that this penalty is lower the higher is $g\left(\lambda\right)$, namely the higher is aggregate giving. Thus, a party's policy-motivated deviation from the mean of district means is more profitable the higher is aggregate giving. Similarly, higher uncertainty Ψ about the mean ideology of party donors reduces the party penalty in terms of lost legislative strength, giving parties a greater incentive to pursue their partisan policy preferences. The following result describes the features of the full equilibrium.

Proposition 4 (Equilibrium Characterization) *There exists a unique symmetric equilibrium of the game. In this equilibrium (i) district candidates adopt symmetric policy po-*

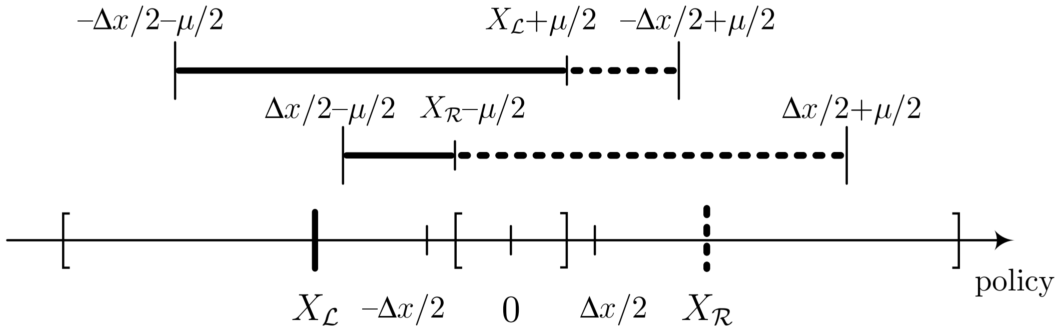


Figure 3: Equilibrium Party and Candidate Strategies.

sitions around the district mean, as characterized in Proposition 2, and (ii) parties adopt symmetric platforms around the mean of district means, $(X_{\mathcal{L}}, X_{\mathcal{R}}) = \left(-\frac{\Delta X}{2}, \frac{\Delta X}{2}\right)$, where $\Delta X = \frac{\Psi}{1+\frac{1/\mu}{g(\lambda)/\Sigma_g}}$. *Parties win equal seat shares. Equilibrium party divergence* ΔX *ranges from zero, when* $g(\lambda) \to 0$, *to* $\frac{\Psi}{1+\Sigma_g/\mu}$, *when* $g(\lambda) \to 1$.

The equilibrium is illustrated in Figure 3. The equilibrium has the property that $X_{\mathcal{L}} \leq -\frac{\Delta x}{2} < \frac{\Delta x}{2} \leq X_{\mathcal{R}}$, which is based on the assumption in equation (7). The equilibrium features the two parties offering symmetric party positions and each party has equal policy weight in the legislature. Thus, each party's equilibrium expected payoff is zero. To establish the existence of this equilibrium one has to rule out all possible deviations. The argument can be illustrated using party $\mathcal{L}$, which by symmetry also applies to party $\mathcal{R}$; the details are in the Appendix. Party $\mathcal{L}$ deviations in the interval $\left(-\infty, -\frac{\Delta x}{2}\right]$ are not profitable because party $\mathcal{L}$'s payoff function in this interval is strictly concave in $X_{\mathcal{L}}$. Party $\mathcal{L}$ deviations in the interval $[0, \infty)$ are not profitable either, because they result at best in a non-negative policy, which gives party $\mathcal{L}$ at most a zero payoff. Party $\mathcal{L}$ deviations in the interval $\left(-\frac{\Delta x}{2}, 0\right)$ yield a negative payoff because as party $\mathcal{L}$ moves its platform $X_{\mathcal{L}}$ beyond $-\frac{\Delta x}{2}$, its policy weight drops faster than a commensurate rightward deviation between $-\frac{\Delta X}{2}$ and $-\frac{\Delta x}{2}$. In the former deviation its seat share shrinks as the other party's stays constant, whereas in the latter deviation its seat share increases as the other party's declines. Thus, party $\mathcal{L}$'s payoff drops faster than if it were to gain seats at the same rate. Uniqueness of the equlibrium follows from the fact that the necessary equilibrium conditions $\frac{\partial}{\partial X_{\mathcal{L}}} U_{\mathcal{L}} = \frac{\partial}{\partial X_{\mathcal{R}}} U_{\mathcal{R}} = 0$ have a single solution, as well as the observation that there are no equilibria with the property that $-\frac{\Delta x}{2} < X_{\mathcal{L}} \leq X_{\mathcal{R}} < \frac{\Delta x}{2}$.

In Figure 3 the equilibrium features districts where parties win with equal probability, located in the middle, as well as districts where one party wins for sure, located toward the extremes. The latter capture what is known as *safe seats*. These districts are partisan strongholds, where the top two serious contenders are from the same party. For instance, California's 12th Congressional District, in the San Francisco area, has been held by the Democratic party without interruption since 1949. A safe seat for the Republicans has been Virginia's 7th Congressional District, in the Richmond area, which has been held by the Republican party since 1971.[23]

Using the necessary equilibrium conditions, one can solve for the equilibrium *party divergence*, namely the difference in party positions:

$$\Delta X \equiv X_{\mathcal{R}} - X_{\mathcal{L}} = \frac{\Psi}{1 + \frac{1/\mu}{g(\lambda)/\Sigma_g}} \tag{15}$$

which is a function of aggregate giving, which in turn depends on income inequality; the degree of uncertainty about mean party donor support; district heterogeneity; and the dispersion of the donor support distribution. The next result states how these factors affect the extent of party divergence.

Proposition 5 (Party Comparative Statics) *Under imperfect aggregate participation,* $0 < g(\lambda) < 1$*, equilibrium party divergence* ΔX *is, all else equal: (i) strictly increasing in income inequality* λ*, (ii) strictly increasing in the uncertainty* Ψ *over the mean ideology of party donor support, and (iii) strictly increasing in across-district ideological heterogeneity* μ*, and strictly decreasing in the dispersion* Σ_g *of party donor support.*

When income inequality increases, it increases aggregate giving, by Proposition 1, making the electoral outcome more dependent on donor support. But, donor support is a polarizing force for parties because the uncertainty it entails about mean donor support reduces the marginal loss of legislative strength for taking partisan positions, reducing the parties' cost from pursuing partisan policy goals. An exogenous increase in the uncertainty Ψ about the mean ideology of party donor support has the same effect on the parties' cost of diverging. Also, higher across-district heterogeneity reduces the rate at which a party loses legislative

[23] This has not always been driven by incumbency advantage, as illustrated by the high-profile defeat in June 2014 of House Majority Leader Eric Cantor (R-VA) by challenger David Brat, a more conservative Republican.

strength as it moves to a more extreme position, since the share of districts lost to the other party for a given move is lower.

4 Model Implications

The model identifies two mechanisms through which the model's exogenous factors affect partisan polarization: party divergence and candidate divergence. These two equilibrium features depend on different, though overlapping, sets of exogenous factors. See the comparative statics predictions in Propositions 3 and 5 above. Therefore, they may move together or separately, depending on which model parameters change.

The two mechanisms naturally allow a decomposition of partisan polarization into two parts, ideological and compositional. *Ideological polarization* is a shift in candidate policy positions without a change in candidate partisan affiliations. In each district, the candidates move their policy positions away from each other, and so do the parties, so that party affiliations remain the same. *Compositional polarization* is a change in candidate partisan affiliations without a shift in candidate policy positions. The candidates maintain their policy positions but some change their party affiliations because the parties have tightened their ideological criteria for party membership. In equilibrium, depending on which model parameters change, the two types of polarization may change in the same direction, or in opposite directions. Figure 4 illustrates these concepts.

The following expression separates equilibrium partisan polarization into its two components.

$$\Delta\mathbb{E}\left(x\right) = \Delta x + \left(\Delta X - \Delta x\right)\left(1 - \frac{\Delta X + \Delta x}{2\mu}\right) \tag{16}$$

Note that compositional polarization, the second term in the above equation, occurs to the extent that party divergence exceeds candidate divergence. The decomposition formula therefore implies the following result.

Proposition 6 (Polarization Mechanisms) *Partisan polarization $\Delta\mathbb{E}\left(x\right)$ is driven by both candidate divergence Δx and party divergence ΔX. Party divergence commensurate with candidate divergence, $\Delta X = \Delta x$, leads to ideological polarization. Party divergence in excess of candidate divergence, $\Delta X > \Delta x$, leads to compositional polarization.*

Party switchers are one manifestation of party divergence that leads to changes in partisan

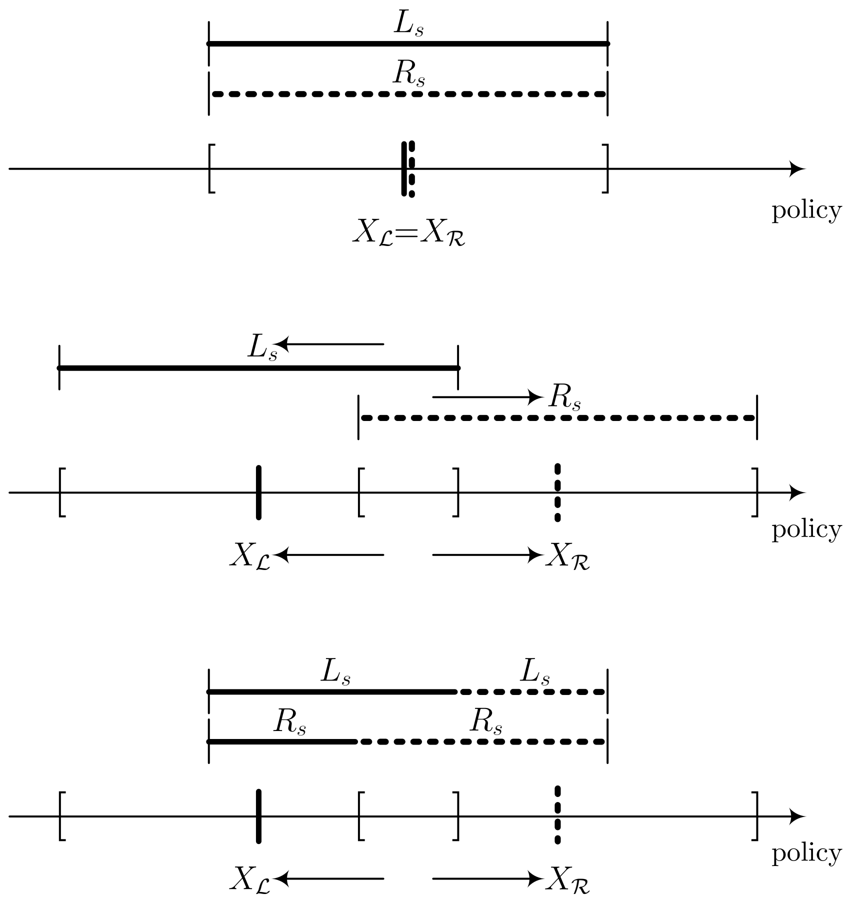


Figure 4: Equilibrium Scenarios. (Top) Zero Polarization - Party and Candidate Convergence. (Middle) Pure Ideological Polarization. (Bottom) Pure Compositional Polarization.

affiliations, and thus to compositional polarization. In 1999 Rep. Michael Forbes (NY) and in 2001 Sen. Jim Jeffords (VT), both liberal Republicans, defected to the Democrats. On the Democratic side, in 1995 Rep. Mike Parker (MS) and in 1994 Sen. Richard Shelby (AL), both conservative Democrats, defected to the GOP.[24]

Apart from partisan polarization, which captures the difference between the means of the parties' policy distributions, other features of the parties' policy distributions display clear trends in the data. See Table 1 below for a summary of these features. Inter-party heterogeneity is the dispersion in the full range of policy positions represented in the legislature. In the model, it can be captured by the policy difference between the rightmost $\mathcal{R}$ party member and the leftmost $\mathcal{L}$ party member. See Figure 3. Intra-party heterogeneity is the dispersion in policy positions within each party, and can be measured as the policy differe-

[24]See Nokken and Poole (2004) for a comprehensive analysis of party defections in the U.S. Congress.

Table 1: Polarization Features

Polarization Feature	Definition	Solution
Partisan Polarization ↗	$\frac{1}{\mu}\left(\int_{k_s\in\mathcal{R}} x_{k_s}ds - \int_{k_s\in\mathcal{L}} x_{k_s}ds\right)$	$\frac{1}{\mu}\left[\frac{(\Delta x)^2}{2} + \frac{\Delta X(2\mu-\Delta X)}{2}\right]$
Inter-Party Heterogeneity ↗	$\max_{k_s\in\mathcal{R}} x_{k_s} - \min_{k_s\in\mathcal{L}} x_{k_s}$	$\mu + \Delta x$
Intra-Party Heterogeneity ↘	$\max_{k_s\in p} x_{k_s} - \min_{k_s\in p} x_{k_s}$	$\mu - \frac{1}{2}(\Delta X - \Delta x)$
Party Overlap ↘	$\max_{k_s\in\mathcal{L}} x_{k_s} - \min_{k_s\in\mathcal{R}} x_{k_s}$	$\mu - \Delta X$
Fraction of Safe Seats ↗	$\frac{1}{\mu}\left(\int_{L_s,R_s\in\mathcal{L}} ds + \int_{L_s,R_s\in\mathcal{R}} ds\right)$	$\frac{1}{\mu}(\Delta X - \Delta x)$

nce between the rightmost and leftmost member within each party. Inter-party heterogeneity has had an upward trend since the post-WWII period, while intra-party heterogeneity has had a downward trend for both parties, in both houses of the U.S. Congress (McCarty, Poole, and Rosenthal 2006, Figures 2.5 and 2.6).[25]

Party overlap can be defined as the common support of the two party memberships' policy positions, namely the policy range between the rightmost $\mathcal{L}$ party member and the leftmost $\mathcal{R}$ party member. A safe seat is a district the same party wins consistently. In the model, the fraction of safe seats can be measured as the fraction of districts where both candidates are affiliated with the same party. In the data, party overlap has decreased (McCarty, Poole, and Rosenthal 2006, Figure 2.9), while the fraction of safe seats has increased (Abramowitz, Alexander, and Gunning 2006, Figure 1).

Note that the different measures of polarization in the first column of Table 1 are observable equilibrium outcomes. In contrast, the mechanisms behind these outcomes, namely party divergence ΔX and candidate divergence Δx, are equilibrium strategies, and may not always be observable. For instance, only one of the top two candidates in a district wins, and the winner's ideology can be estimated based on his/her voting record in Congress. The losing candidate's ideology - assuming he/she never served before - is unobservable.

For a factor of the model to explain all the empirical polarization features of Table 1, it has to simultaneously exert the following four effects on the equilibrium:

(i) candidate divergence Δx increases.

(ii) party divergence ΔX increases.

(iii) the party-candidate divergence differential $\Delta X - \Delta x$ is positive.

[25] Diermeier, Prato, and Vlaicu (2020) study how inter-party heterogeneity and intra-party homogeneity affect legislative policy through the adoption of partisan procedural rules.

(iv) the party-candidate divergence differential $\Delta X - \Delta x$ increases.

This is apparent from the last column of Table 1. For part (iii), recall that intra-party heterogeneity cannot exceed district heterogeneity μ.

What parameters of the model are consistent with the implications (i)-(iv) coming from the data? Apart from income inequality λ, no other model parameter by itself can simultaneously increase both candidate and party divergence. If the model's parameters satisfy the first part of equation (7) with strict inequality, then property (iii) holds. Finally, if the following condition holds, then income inequality can also lead to property (iv).

$$\frac{\psi\left[v\left(\lambda\right)g'\left(\lambda\right)-g\left(\lambda\right)v'\left(\lambda\right)\right]}{\left[1+\frac{v(\lambda)/\sigma_v}{g(\lambda)/\sigma_g}\right]^2}\frac{\sigma_g}{\sigma_v}<\frac{\Psi g'\left(\lambda\right)}{\left[1+\frac{1/\mu}{g(\lambda)/\Sigma_g}\right]^2}\frac{\Sigma_g}{\mu} \tag{17}$$

The following result summarizes the effects of income inequality on the various features of polarization.

Proposition 7 (Inequality Effects) *An increase in income inequality λ increases inter-party ideological heterogeneity and reduces party overlap. Under condition (17), an increase in income inequality λ reduces intra-party ideological heterogeneity and increases the fraction of safe seats.*

Relating back to the broader literature on elite polarization, three prominent electoral explanations have been: within-district citizen polarization, gerrymandering, and the "big sort." Since the model nests the factors that underlie these alternative explanations it may be instructive to discuss the implications of changing these factors by themselves, while keeping income inequality constant. The following discussion is based on the comparative statics derived in Propositions 3 and 5.

An increase in within-district polarization can be captured by an increase in within-district ideological heterogeneity σ_v. This increases candidate divergence, but has no effect on party divergence. Thus, partisan polarization increases, but only through its ideological polarization component. Gerrymandering may be modeled as an increase in across-district heterogeneity μ. That leads to an increase in party divergence; however, candidate divergence should be unaffected. Thus, partisan polarization increases, but only through its compositional polarization component.[26] The "big sort" argument (Bishop 2008) is that po-

[26] A caveat of this argument is that a partisan gerrymander may not always result in increased district

larization has been driven by electoral districts becoming more internally homogenous and more distinct from each other. In the model this change can be captured by a decrease in within-district ideological heterogeneity σ_v and a concomitant increase in across-district heterogeneity μ. Proposition 3 implies that candidate divergence decreases, while Proposition 5 implies that party divergence increases. Thus, the net effect on partisan polarization $\Delta\mathbb{E}(x)$ can go either way. See equation (16). To generate all the stylized facts noted above both within- and across-district heterogeneity would have to simultaneously increase, a scenario that none of these explanations considers.

Proposition 7 also has implications for the empirical analysis of gerrymandering. The standard grerrymandering argument is that the artificial creation of safe districts by state legislatures has reduced the aggregate responsiveness of national representatives to the preferences of their constituents. Krasa and Polborn (2014b), however, argue that the empirical approach has lead to incorrect inferences about the impact of gerrymandering because it did not take into account strategic spillover effects across districts. Proposition 7 provides another cautionary note to the standard empirical approach to measuring gerrymandering effects. It implies that the share of safe seats typically associated with gerrymandering can also be a by-product, rather than the cause, of high partisan polarization. In the model, parties strategically "retreat" from some districts in order to attract more extreme candidates at the opposite end of the ideological spectrum. This strategic party radicalization is driven by changes in citizen participation patterns that have their source in the increased dispersion of the income distribution.

5 Conclusion

The close association between economic inequality and political polarization documented in the U.S. data has been explained by the previous literature by intensified class conflict or by a political wealth bias. This paper provides an alternative perspective on inequality and polarization. It provides a model where inequality affects polarization by changing the relative prevalence of the two main types of political participation: voting and political giving. By changing participation patterns, inequality alters the strategic logic of political

heterogeneity, as the majority party creates many moderate districts to dilute the power of the minority party. The gerrymandering argument is limited, however, also because it cannot account for the polarization seen in the U.S. Senate, where constituencies (state borders) are fixed.

competition, from an emphasis on centrist voters to an emphasis on partisan policy goals. The key message of this paper, therefore, is that economic inequality may lead to political polarization not simply because it changes the electorate's policy preferences or gives a political advantage to high-income citizens, but because of increased dependence of elections on giving relative to voting. This reduces the electoral penalty for partisan policies and leads to both candidate and party divergence. The model implies that parties have diverged more than candidates, and that the gap between party and candidate divergence has increased with income inequality. These two mechanisms create both ideological and compositional polarization in the policy positions of members of the two parties. Changes in income inequality can also explain other interesting features of the U.S. polarization experience, such as intra-party homogeneity and the rise in safe seats.

The framework proposed here can be extended to study related questions. The model implies that partisan polarization can be mitigated by reducing the convexity of individual giving in income, e.g., by capping contributions from high-income citizens, or reducing the concavity of individual voting in income, e.g., by incentivizing voting by low-income citizens. Modeling the microfoundations of political giving would lead to a setting more amenable to a welfare analysis of campaign finance or electoral reform, such as contributions limits, voting subsidies, or absentee fines.

Candidates' motivations to affiliate with parties may not be purely policy-based, as assumed here, but also based on electoral considerations, e.g., as in Snyder and Ting (2002). Exploring the affiliation decision in greater depth would provide further insight into the compositional component of partisan polarization. One could also relax the assumption that candidates care only about their policy position, and allow them to also value national policy goals, as in Krasa and Polborn (2014b). A more challenging direction is to introduce a feedback loop from polarization to participation that would permit an analysis of turnout dynamics, in a manner similar to Feddersen and Gul (2015) who study the two-way relationship between inequality and polarization when policies affect donor preferences. On the one hand, high polarization may energize previously apathetic voters that did not see significant differences between candidates; on the other hand, high polarization may alienate centrist voters that do not feel well represented by either candidate.

Appendix

Proof of Proposition 1. Consider two income distributions $Y_{\lambda_1}(y)$ and $Y_{\lambda_2}(y)$ with inequality levels λ_1, λ_2 and equal means: $\int_0^\infty y dY_{\lambda_1}(y) = \int_0^\infty y dY_{\lambda_2}(y) = 1$. It is known that under these conditions, $\lambda_1 < \lambda_2$ if and only if $Y_{\lambda_1}(y)$ second-order stochastically dominates $Y_{\lambda_2}(y)$. Then, by SOSD, because $\nu(y)$ is strictly concave in income y, it follows that $v(\lambda_1) = \int_0^\infty \nu(y)\, dY_{\lambda_1}(y) > \int_0^\infty \nu(y)\, dY_{\lambda_2}(y) = v(\lambda_2)$, implying that $\frac{dv(\lambda)}{d\lambda} < 0$. Because $\gamma(y)$ is strictly convex in income y, we have $g(\lambda_1) = \int_0^\infty \gamma(y)\, dY_{\lambda_1}(y) < \int_0^\infty \gamma(y)\, dY_{\lambda_2}(y) = g(\lambda_2)$, implying that $\frac{dg(\lambda)}{d\lambda} > 0$. ■

Proof of Proposition 2. Solve for an L candidate's winning probability:

$$\begin{aligned}
w_{L_s}(\overline{x}_s, \lambda) &= \mathbb{P}\left\{ \frac{[Z_s(\overline{x}_s)]^{\frac{v(\lambda)}{v(\lambda)+g(\lambda)}} \left[D_s\left(\overline{x}_s|\tilde{\psi}_s\right)\right]^{\frac{g(\lambda)}{v(\lambda)+g(\lambda)}}}{[1 - Z_s(\overline{x}_s)]^{\frac{v(\lambda)}{v(\lambda)+g(\lambda)}} \left[1 - D_s\left(\overline{x}_s|\tilde{\psi}_s\right)\right]^{\frac{g(\lambda)}{v(\lambda)+g(\lambda)}}} \geq 1 \right\} \qquad (18)\\
&= \mathbb{P}\left\{ \frac{\left[\frac{1}{1+e^{-\frac{\overline{x}_s - s}{\sigma_v}}}\right]^{\frac{v(\lambda)}{v(\lambda)+g(\lambda)}} \left[\frac{1}{1+e^{-\frac{\overline{x}_s - s - \tilde{\psi}_s}{\sigma_g}}}\right]^{\frac{g(\lambda)}{v(\lambda)+g(\lambda)}}}{\left[1 - \frac{1}{1+e^{-\frac{\overline{x}_s - s}{\sigma_v}}}\right]^{\frac{v(\lambda)}{v(\lambda)+g(\lambda)}} \left[1 - \frac{1}{1+e^{-\frac{\overline{x}_s - s - \tilde{\psi}_s}{\sigma_g}}}\right]^{\frac{g(\lambda)}{v(\lambda)+g(\lambda)}}} \geq 1 \right\}\\
&= \mathbb{P}\left\{ e^{\frac{\overline{x}_s - s}{\sigma_v}\frac{v(\lambda)}{v(\lambda)+g(\lambda)}} e^{\frac{\overline{x}_s - s - \tilde{\psi}_s}{\sigma_g}\frac{g(\lambda)}{v(\lambda)+g(\lambda)}} \geq 1 \right\}\\
&= \mathbb{P}\left\{ \left[\frac{v(\lambda)}{\sigma_v} + \frac{g(\lambda)}{\sigma_g}\right](\overline{x}_s - s) \geq \tilde{\psi}_s \frac{g(\lambda)}{\sigma_g} \right\} = \mathbb{P}\left\{ \tilde{\psi}_s \leq \left[1 + \frac{v(\lambda)/\sigma_v}{g(\lambda)/\sigma_g}\right](\overline{x}_s - s) \right\}\\
&= \frac{1}{2} + \frac{1}{\psi}\left[1 + \frac{v(\lambda)/\sigma_v}{g(\lambda)/\sigma_g}\right]\left(\frac{x_{L_s} + x_{R_s}}{2} - s\right) \qquad (19)
\end{aligned}$$

and $w_{R_s}(\overline{x}_s, \lambda) = 1 - w_{L_s}(\overline{x}_s, \lambda)$.

The necessary conditions for a candidate equilibrium are, using $w_{L_s} + w_{R_s} = 1$:

$$\begin{aligned}
\frac{\partial}{\partial x_{L_s}} u_{L_s} &= -w_{L_s} + (x_{R_s} - x_{L_s}) \frac{\partial}{\partial x_{L_s}} w_{L_s} = 0 \qquad (20)\\
\frac{\partial}{\partial x_{R_s}} u_{R_s} &= (1 - w_{L_s}) - (x_{R_s} - x_{L_s}) \frac{\partial}{\partial x_{R_s}} w_{L_s} = 0 \qquad (21)
\end{aligned}$$

and adding these two first-order conditions, and observing that $\frac{\partial}{\partial x_{L_s}} w_{L_s} = \frac{\partial}{\partial x_{R_s}} w_{L_s} = \frac{1}{2\psi}\left[1 + \frac{v(\lambda)/\sigma_v}{g(\lambda)/\sigma_g}\right]$ from equation (19), yields $w_{L_s} = w_{R_s} = \frac{1}{2}$. Using this in the first-order conditions, one can solve for the candidate equilibrium strategies $(x_{L_s}, x_{R_s}) = \left(s - \frac{\Delta x}{2}, s + \frac{\Delta x}{2}\right)$ where candidate divergence is:

$$\Delta x \equiv x_{R_s} - x_{L_s} = \frac{w_{L_s}}{\frac{\partial}{\partial x_{L_s}} w_{L_s}} = \frac{\psi}{1 + \frac{v(\lambda)/\sigma_v}{g(\lambda)/\sigma_g}} \tag{22}$$

To establish that there are no global profitable deviations the following properties suffice: (i) $u_{L_s}(x_{L_s}, x_{R_s})$ is strictly concave in x_{L_s} on $(-\infty, x_{R_s}]$, which implies that the x_{L_s} that solves equations (20)-(21) is a global maximizer of $u_{L_s}(x_{L_s}, x_{R_s})$ on $(-\infty, x_{R_s}]$, and (ii) a candidate L_s deviation from the equilibrium x_{L_s} to some $x'_{L_s} > x_{R_s}$ strictly lowers L_s's payoff. The argument for candidate R_s is analogous. For part (i) note that $\frac{\partial^2}{\partial x_{L_s}^2} u_{L_s}(x_{L_s}, x_{R_s}) = -2\frac{\partial}{\partial x_{L_s}} w_{L_s} + (x_{R_s} - x_{L_s}) \frac{\partial^2}{\partial x_{L_s}^2} w_{L_s}$ and since by equation (19) we have that w_{L_s} is linear and strictly increasing in x_{L_s}, it follows that $\frac{\partial^2}{\partial x_{L_s}^2} u_{L_s}(x_{L_s}, x_{R_s}) = -2\frac{\partial}{\partial x_{L_s}} w_{L_s} < 0$. For part (ii) note that the policy lottery induced by the deviation is first-order stochastically dominated for L_s since $u_{L_s}(x'_{L_s}, x_{R_s}) < u_{L_s}(x_{R_s}, x_{R_s}) = -x_{R_s} \leq -\frac{x_{L_s}+x_{R_s}}{2} = u_{L_s}(x_{L_s}, x_{R_s})$. ■

Proof of Proposition 3. A candidate's marginal electoral penalty for moving away from the district mean ideology s, namely $\frac{1}{2\psi}\left[1 + \frac{v(\lambda)/\sigma_v}{g(\lambda)/\sigma_g}\right]$ is lower the lower the ratio $v(\lambda)/g(\lambda)$, namely the more important is giving relative to voting. Thus, a candidate policy-motivated deviation from the district mean ideology is more profitable the lower the ratio $v(\lambda)/g(\lambda)$. Taking the difference between the two first-order conditions (20)-(21) and solving for $x_{R_s} - x_{L_s}$ yields $\Delta x = \frac{1}{\frac{\partial}{\partial x_{L_s}} w_{L_s} + \frac{\partial}{\partial x_{R_s}} w_{L_s}} = \frac{\psi}{1+\frac{v(\lambda)/\sigma_v}{g(\lambda)/\sigma_g}}$. Then, using $v'(\lambda) < 0$ and $g'(\lambda) > 0$ from Proposition 1 gives that $\frac{\partial}{\partial\lambda}\Delta x > 0$. The comparative statics with respect to ψ, σ_v, σ_g follow immediately from the expression for Δx. ■

Proof of Proposition 4. Parties anticipate that candidates will play the strategies characterized in Proposition 2. Given candidate strategies, parties calculate their seat shares and expected donations, which together will determine their policy weights. To show equilibrium existence and uniqueness, the proof consists of the following steps. (i) Necessary conditions for a party equilibrium in the case $X_{\mathcal{L}} \leq -\frac{\Delta x}{2} \leq \frac{\Delta x}{2} \leq X_{\mathcal{R}}$. (ii) Ruling out global deviations. (iii) Ruling out a party equilibrium with $-\frac{\Delta x}{2} < X_{\mathcal{L}} \leq X_{\mathcal{R}} < \frac{\Delta x}{2}$.

(i) Consider an equilibrium with the property $X_{\mathcal{L}} \leq -\frac{\Delta x}{2} \leq \frac{\Delta x}{2} \leq X_{\mathcal{R}}$. To solve for the parties' policy weights, it is necessary to derive the parties' seat shares. See equation (12) and Figure 3.

$$
\begin{aligned}
h_{\mathcal{L}} &= \frac{1}{\mu}\left\{\frac{1}{2}\left[\left(X_{\mathcal{L}}+\frac{\mu}{2}\right)-\left(-\frac{\Delta x}{2}-\frac{\mu}{2}\right)\right]+\frac{1}{2}\left[\left(X_{\mathcal{R}}-\frac{\mu}{2}\right)-\left(\frac{\Delta x}{2}-\frac{\mu}{2}\right)\right]\right\} \\
&= \frac{1}{\mu}\left(\frac{\mu}{2}+\frac{X_{\mathcal{L}}+X_{\mathcal{R}}}{2}\right) && (23) \\
h_{\mathcal{R}} &= \frac{1}{\mu}\left\{\frac{1}{2}\left[\left(\frac{\Delta x}{2}+\frac{\mu}{2}\right)-\left(X_{\mathcal{R}}-\frac{\mu}{2}\right)\right]+\frac{1}{2}\left[\left(-\frac{\Delta x}{2}+\frac{\mu}{2}\right)-\left(X_{\mathcal{L}}+\frac{\mu}{2}\right)\right]\right\} \\
&= \frac{1}{\mu}\left(\frac{\mu}{2}-\frac{X_{\mathcal{L}}+X_{\mathcal{R}}}{2}\right) && (24)
\end{aligned}
$$

and thus $h_{\mathcal{L}}-h_{\mathcal{R}}=\frac{1}{\mu}\left(X_{\mathcal{L}}+X_{\mathcal{R}}\right)=\frac{2\bar{X}}{\mu}$. Policy weights can then be calculated as follows.

$$
\begin{aligned}
W_{\mathcal{L}}\left(\bar{X},\lambda\right) &= \mathbb{P}\left\{\frac{\left[Z\left(\frac{2\bar{X}}{\mu}\right)\right]^{\frac{1}{1+g(\lambda)}}\left[D\left(\bar{X}|\tilde{\Psi}\right)\right]^{\frac{g(\lambda)}{1+g(\lambda)}}}{\left[1-Z\left(\frac{2\bar{X}}{\mu}\right)\right]^{\frac{1}{1+g(\lambda)}}\left[1-D\left(\bar{X}|\tilde{\Psi}\right)\right]^{\frac{g(\lambda)}{1+g(\lambda)}}}\geq 1\right\} && (25) \\
&= \mathbb{P}\left\{\frac{\left[\frac{1}{1+e^{-\frac{\bar{X}}{\mu}}}\right]^{\frac{1}{1+g(\lambda)}}\left[\frac{1}{1+e^{-\frac{\bar{X}-\tilde{\Psi}}{\Sigma_g}}}\right]^{\frac{g(\lambda)}{1+g(\lambda)}}}{\left[1-\frac{1}{1+e^{-\frac{\bar{X}}{\mu}}}\right]^{\frac{1}{1+g(\lambda)}}\left[1-\frac{1}{1+e^{-\frac{\bar{X}-\tilde{\Psi}}{\Sigma_g}}}\right]^{\frac{g(\lambda)}{1+g(\lambda)}}}\geq 1\right\} \\
&= \mathbb{P}\left\{e^{\frac{\bar{X}}{\mu}\frac{1}{1+g(\lambda)}}e^{\frac{\bar{X}-\tilde{\Psi}}{\Sigma_g}\frac{g(\lambda)}{1+g(\lambda)}}\geq 1\right\}=\mathbb{P}\left\{\frac{\bar{X}}{\mu}+\frac{\bar{X}-\tilde{\Psi}}{\Sigma_g}g\left(\lambda\right)\geq 0\right\} \\
&= \mathbb{P}\left\{\frac{\bar{X}}{\mu}+\frac{\bar{X}}{\Sigma_g}g\left(\lambda\right)\geq g\left(\lambda\right)\frac{\tilde{\Psi}}{\Sigma_g}\right\}=\mathbb{P}\left\{\tilde{\Psi}\leq\left[1+\frac{1/\mu}{g\left(\lambda\right)/\Sigma_g}\right]\bar{X}\right\} \\
&= \frac{1}{2}+\frac{1}{\Psi}\left[1+\frac{1/\mu}{g\left(\lambda\right)/\Sigma_g}\right]\frac{X_{\mathcal{L}}+X_{\mathcal{R}}}{2} && (26)
\end{aligned}
$$

and $W_{\mathcal{R}}\left(\bar{X},\lambda\right)=1-W_{\mathcal{L}}\left(\bar{X},\lambda\right)$. The necessary conditions for a party equilibrium are, using $W_{\mathcal{L}}+W_{\mathcal{R}}=1$:

$$
\begin{aligned}
\frac{\partial}{\partial X_{\mathcal{L}}}U_{\mathcal{L}} &= -W_{\mathcal{L}}+\left(X_{\mathcal{R}}-X_{\mathcal{L}}\right)\frac{\partial}{\partial X_{\mathcal{L}}}W_{\mathcal{L}}=0 && (27) \\
\frac{\partial}{\partial X_{\mathcal{R}}}U_{\mathcal{R}} &= \left(1-W_{\mathcal{L}}\right)-\left(X_{\mathcal{R}}-X_{\mathcal{L}}\right)\frac{\partial}{\partial X_{\mathcal{R}}}W_{\mathcal{L}}=0 && (28)
\end{aligned}
$$

and adding together these two first-order conditions, and observing that $\frac{\partial}{\partial X_{\mathcal{L}}}W_{\mathcal{L}}=\frac{\partial}{\partial X_{\mathcal{R}}}W_{\mathcal{L}}=\frac{1}{2\Psi}\left[1+\frac{1/\mu}{g(\lambda)/\Sigma_g}\right]$ from equation (26), yields $W_{\mathcal{L}}=W_{\mathcal{L}}=\frac{1}{2}$. Using this in the first-order conditions, one can solve for the party equilibrium strategies $(X_{\mathcal{L}}, X_{\mathcal{R}}) = \left(-\frac{\Delta X}{2}, \frac{\Delta X}{2}\right)$ where:

$$\Delta X \equiv X_{\mathcal{R}} - X_{\mathcal{L}} = \frac{W_{\mathcal{L}}}{\frac{\partial}{\partial X_{\mathcal{L}}} W_{\mathcal{L}}} = \frac{\Psi}{1 + \frac{1/\mu}{g(\lambda)/\Sigma_g}} \tag{29}$$

(ii) To establish that the party strategies above constitute an equilibrium, not only local but also global deviations have to be unprofitable. The argument is made for party $\mathcal{L}$ only, since by symmetry it also applies to party $\mathcal{R}$. First note that parties' payoffs in the conjectured equilibrium are zero by symmetry. Party $\mathcal{L}$ deviations in the interval $\left(-\infty, -\frac{\Delta x}{2}\right]$ are not profitable because party $\mathcal{L}$'s payoff function in this interval is strictly concave in $X_{\mathcal{L}}$. To see this, note that $\frac{\partial^2}{\partial X_{\mathcal{L}}^2} U_{\mathcal{L}} = -2 \frac{\partial}{\partial X_{\mathcal{L}}} W_{\mathcal{L}} + (X_{\mathcal{R}} - X_{\mathcal{L}}) \frac{\partial^2}{\partial X_{\mathcal{L}}^2} W_{\mathcal{L}} = -2 \frac{\partial}{\partial X_{\mathcal{L}}} W_{\mathcal{L}} < 0$ because $W_{\mathcal{L}}$ is strictly increasing and linear in $X_{\mathcal{L}}$. See equation (26). Party $\mathcal{L}$ deviations in the interval $[0, \infty)$ are not profitable because they produce at best a non-negative policy, which gives party $\mathcal{L}$ at most a zero payoff. Party $\mathcal{L}$ deviations in the interval $\left(-\frac{\Delta x}{2}, 0\right)$ yield a negative payoff. To see this, note that as party $\mathcal{L}$ moves its position $X_{\mathcal{L}}$ beyond $-\frac{\Delta x}{2}$, its policy weight $W_{\mathcal{L}}$ drops faster than a commensurate rightward deviation between $-\frac{\Delta X}{2}$ and $-\frac{\Delta x}{2}$, because in the former deviation its seat share shrinks as the other party's stays constant, whereas in the latter deviation its seat share increases as the other party's declines. See Figure 3. Thus, party $\mathcal{L}$'s payoff drops faster than if it were to gain seats at the same rate.

(iii) To establish uniqueness, an equilibrium with the property $-\frac{\Delta x}{2} < X_{\mathcal{L}} \leq X_{\mathcal{R}} < \frac{\Delta x}{2}$ has to be ruled out. Suppose parties pick positions with these properties. To solve for the parties' policy weights, it is necessary to first derive the parties' seat shares:

$$h_{\mathcal{L}} = \frac{1}{\mu}\left\{\frac{1}{2}\left[\left(-\frac{\Delta x}{2} + \frac{\mu}{2}\right) - \left(X_{\mathcal{L}} - \frac{\mu}{2}\right)\right]\right\} = \frac{1}{2\mu}\left(\mu - \frac{\Delta x}{2} - X_{\mathcal{L}}\right) \tag{30}$$

$$h_{\mathcal{R}} = \frac{1}{\mu}\left\{\frac{1}{2}\left[\left(X_{\mathcal{R}} + \frac{\mu}{2}\right) - \left(\frac{\Delta x}{2} - \frac{\mu}{2}\right)\right]\right\} = \frac{1}{2\mu}\left(\mu - \frac{\Delta x}{2} + X_{\mathcal{R}}\right) \tag{31}$$

and thus $h_{\mathcal{L}} - h_{\mathcal{R}} = -\frac{1}{2\mu}(X_{\mathcal{L}} + X_{\mathcal{R}}) = -\frac{\bar{X}}{\mu}$. Following a computation as in equation (26), policy weights are as follows:

$$W_{\mathcal{L}}\left(\bar{X}, \lambda\right) = \frac{1}{2} + \frac{1}{\Psi}\left[1 - \frac{1/2\mu}{g(\lambda)/\Sigma_g}\right] \frac{X_{\mathcal{L}} + X_{\mathcal{R}}}{2} \tag{32}$$

and $W_{\mathcal{R}}\left(\bar{X}, \lambda\right) = 1 - W_{\mathcal{L}}\left(\bar{X}, \lambda\right)$. Solving the necessary equilibrium conditions in equations

(27)-(28), gives the party equilibrium strategies $(X_{\mathcal{L}}, X_{\mathcal{R}}) = \left(-\frac{\Delta X}{2}, \frac{\Delta X}{2}\right)$ where:

$$\Delta X \equiv X_{\mathcal{R}} - X_{\mathcal{L}} = \frac{W_{\mathcal{L}}}{\frac{\partial}{\partial X_{\mathcal{L}}} W_{\mathcal{L}}} = \frac{\Psi}{1 - \frac{1/2\mu}{g(\lambda)/\Sigma_g}} \tag{33}$$

However, since $\frac{\Psi}{1-\frac{1/2\mu}{g(\lambda)/\Sigma_g}} > \frac{\Psi}{1+\frac{1/\mu}{g(\lambda)/\Sigma_g}}$, this larger party divergence not consistent with the restriction $-\frac{\Delta x}{2} < X_{\mathcal{L}} \leq X_{\mathcal{R}} < \frac{\Delta x}{2}$, which implies a smaller party divergence, thus ruling out an equilibrium in this range. ■

Proof of Proposition 5. A party's marginal electoral penalty for moving away from the district mean, namely $\frac{1}{2\Psi}\left[1 + \frac{1/\mu}{g(\lambda)/\Sigma_g}\right]$ is lower the higher is aggregate giving $g(\lambda)$. Thus, a party's policy-motivated deviation from the mean of district means is more profitable the higher is $g(\lambda)$. Taking the difference between the two first-order conditions (27)-(28) and solving for $X_{\mathcal{R}} - X_{\mathcal{L}}$ yields $\Delta X = \frac{1}{\frac{\partial}{\partial X_{\mathcal{L}}} W_{\mathcal{L}} + \frac{\partial}{\partial X_{\mathcal{R}}} W_{\mathcal{L}}} = \frac{\Psi}{1+\frac{1/\mu}{g(\lambda)/\Sigma_g}}$. Then, using $g'(\lambda) > 0$ from Proposition 1 gives that $\frac{\partial}{\partial \lambda}\Delta X > 0$. The comparative statics with respect to Ψ, μ, Σ_g follow immediately from the expression for ΔX. ■

Proof of Proposition 6. Equilibrium partisan polarization can be expressed as a function of candidate divergence and party divergence.

$$\begin{aligned}
\Delta \mathbb{E}(x) &= \frac{1}{\mu}\left(\int_{k_s \in \mathcal{R}} x_{k_s} ds - \int_{k_s \in \mathcal{L}} x_{k_s} ds\right) = \frac{2}{\mu}\int_{k_s \in \mathcal{R}} x_{k_s} ds \\
&= \frac{2}{\mu}\left[\left(\frac{\mu}{2} + \frac{\Delta x}{2}\right) - \left(X_{\mathcal{R}} - \frac{\mu}{2}\right)\right]\left[\left(X_{\mathcal{R}} - \frac{\mu}{2}\right) + \left(\frac{\mu}{2} + \frac{\Delta x}{2}\right)\right]/2 \\
&\quad + \frac{2}{\mu}\left[\left(\frac{\mu}{2} - \frac{\Delta x}{2}\right) - \left(X_{\mathcal{L}} + \frac{\mu}{2}\right)\right]\left[\left(X_{\mathcal{L}} + \frac{\mu}{2}\right) + \left(\frac{\mu}{2} - \frac{\Delta x}{2}\right)\right]/2 \\
&= \frac{1}{\mu}\left\{\left[\left(\frac{\mu}{2} + \frac{\Delta x}{2}\right)^2 - \left(X_{\mathcal{R}} - \frac{\mu}{2}\right)^2\right] + \left[\left(\frac{\mu}{2} - \frac{\Delta x}{2}\right)^2 - \left(X_{\mathcal{L}} + \frac{\mu}{2}\right)^2\right]\right\} \\
&= \frac{1}{\mu}\left[\frac{\mu \Delta x}{2} + \frac{(\Delta x)^2}{4} - X_{\mathcal{R}}^2 + X_{\mathcal{R}}\mu - \frac{\mu \Delta x}{2} + \frac{(\Delta x)^2}{4} - X_{\mathcal{L}}^2 - X_{\mathcal{L}}\mu\right] \\
&= \frac{1}{\mu}\left[\frac{(\Delta x)^2}{2} - X_{\mathcal{R}}^2 + X_{\mathcal{R}}\mu - X_{\mathcal{L}}^2 - X_{\mathcal{L}}\mu\right] = \frac{1}{\mu}\left[\frac{(\Delta x)^2}{2} - \frac{(\Delta X)^2}{2} + \mu \Delta X\right] \\
&= \frac{1}{\mu}\left[\frac{(\Delta x)^2}{2} + \frac{\Delta X(2\mu - \Delta X)}{2}\right] = \Delta x + (\Delta X - \Delta x)\left(1 - \frac{\Delta X + \Delta x}{2\mu}\right)
\end{aligned} \tag{34}$$

Ideological polarization is Δx and compositional polarization is $(\Delta X - \Delta x)\left(1 - \frac{\Delta X + \Delta x}{2\mu}\right)$, which is positive if and only if $\Delta X > \Delta x$. If $\Delta X = \Delta x$, then $\Delta \mathbb{E}(x) = \Delta x$ is pure ideological polarization. See Figure 4, middle panel. If $\Delta x = 0$, then $\Delta \mathbb{E}(x) = \frac{\Delta X(2\mu - \Delta X)}{2\mu}$ is pure compositional polarization. See Figure 4, bottom panel. ■

Proof of Proposition 7. The derivations of the different features of equilibrium polarization are as follows. See also Figure 3. Inter-party heterogeneity is the policy difference between the rightmost $\mathcal{R}$ party member and the leftmost $\mathcal{L}$ party member:

$$\max_{k_s \in \mathcal{R}} x_{k_s} - \min_{k_s \in \mathcal{L}} x_{k_s} = \left(\frac{\Delta x}{2} + \frac{\mu}{2}\right) - \left(-\frac{\Delta x}{2} - \frac{\mu}{2}\right) = \mu + \Delta x \tag{35}$$

Intra-party heterogeneity is the policy difference between the rightmost and leftmost member within each party, and by equilibrium symmetry, is the same in both parties:

$$\begin{aligned} \max_{k_s \in \mathcal{R}} x_{k_s} - \min_{k_s \in \mathcal{R}} x_{k_s} &= \left(\frac{\Delta x}{2} + \frac{\mu}{2}\right) - \left(X_{\mathcal{R}} - \frac{\mu}{2}\right) = \mu - \frac{1}{2}(\Delta X - \Delta x) && (36) \\ \max_{k_s \in \mathcal{L}} x_{k_s} - \min_{k_s \in \mathcal{L}} x_{k_s} &= \left(X_{\mathcal{L}} + \frac{\mu}{2}\right) - \left(-\frac{\Delta x}{2} - \frac{\mu}{2}\right) = \mu - \frac{1}{2}(\Delta X - \Delta x) && (37) \end{aligned}$$

as $X_{\mathcal{R}} = -X_{\mathcal{L}} = \frac{\Delta X}{2}$ in equilibrium. Party overlap is the policy difference between the rightmost $\mathcal{L}$ party member and the leftmost $\mathcal{R}$ party member:

$$\max_{k_s \in \mathcal{L}} x_{k_s} - \min_{k_s \in \mathcal{R}} x_{k_s} = \left(X_{\mathcal{L}} + \frac{\mu}{2}\right) - \left(X_{\mathcal{R}} - \frac{\mu}{2}\right) = \mu - \Delta X \tag{38}$$

The fraction of safe seats is the fraction of districts where both candidates are affiliated with the same party:

$$\begin{aligned} \frac{1}{\mu}\left(\int_{L_s, R_s \in \mathcal{L}} ds + \int_{L_s, R_s \in \mathcal{R}} ds\right) &= \frac{1}{\mu}\left\{\left[\left(X_{\mathcal{R}} - \frac{\mu}{2}\right) - \left(\frac{\Delta x}{2} - \frac{\mu}{2}\right)\right]\right. \\ &\quad \left. + \left[\left(-\frac{\Delta x}{2} + \frac{\mu}{2}\right) - \left(X_{\mathcal{L}} + \frac{\mu}{2}\right)\right]\right\} \\ &= \frac{1}{\mu}\left[(X_{\mathcal{R}} - X_{\mathcal{L}}) - \Delta x\right] = \frac{1}{\mu}(\Delta X - \Delta x) && (39) \end{aligned}$$

According to Propositions 3 and 5, both Δx and ΔX are strictly increasing in inequality λ. Condition (17) implies that $\frac{\partial}{\partial \lambda}\Delta X > \frac{\partial}{\partial \lambda}\Delta x$, which means that $\Delta X - \Delta x$ is strictly increasing in inequality. ■

References

Abramowitz, A.I. (2010) *The Disappearing Center: Engaged Citizens, Polarization, and American Democracy*, New Haven, CT: Yale University Press.

Abramowitz, A.I., B. Alexander, and M. Gunning (2006) "Incumbency, Redistricting, and the Decline of Competition in U.S. House Elections," *Journal of Politics* 68(1): 75-88.

Ansolabehere, S., J.M. de Figueiredo, J.M. Snyder Jr. (2003) "Why Is There so Little Money in U.S. Politics?" *Journal of Economic Perspectives* 17(1): 105-130.

Ansolabehere, S. and E. Hersh (2012) "Validation: What Big Data Reveal About Survey Misreporting and the Real Electorate" *Political Analysis* 20(4): 437-459.

Austen-Smith, D. (1984) "Two-Party Competition with Many Constituencies," *Mathematical Social Sciences* 7(2): 177-198.

Azzimonti, M. (2015) "Partisan Conflict and Private Investment," NBER Working Paper 21273.

Bartels, L.M. (2008) *Unequal Democracy: The Political Economy of the New Gilded Age*, Princeton, NJ: Princeton University Press.

Bishop, B. (2008) *The Big Sort: Why the Clustering of Like-minded America is Tearing Us Apart*, Houghton Mifflin Harcourt.

Bonica, A. (2014) "Mapping the Ideological Marketplace," *American Journal of Political Science* 58: 367–386.

Bonica, A., N. McCarty, K.T. Poole, and H. Rosenthal (2013) "Why Hasn't Democracy Slowed Rising Inequality?" *Journal of Economic Perspectives* 27(3): 103-124.

Brody, R.A. (1978) "The Puzzle of Political Participation in America," in A. King (ed.) *The New American Political System*, Washington, DC: AEI.

Callander, S. (2005) "Electoral Competition in Heterogenous Districts," *Journal of Political Economy* 113(5): 1116-1145.

Callander, S. and C. Wilson (2007) "Turnout, Polarization, and Duverger's Law," *Journal of Politics* 69(4): 1047–1056.

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

Campante, F. (2011) "Redistribution in a Model of Voting and Campaign Contributions" *Journal of Public Economics* 95(7-8): 646-656.

Diermeier, D. and R. Vlaicu (2011) "Parties, Coalitions, and the Internal Organization of Legislatures," *American Political Science Review* 105(2): 359-380.

Diermeier, D., C. Prato, and R. Vlaicu (2015) "Procedural Choice in Majoritarian Organizations," *American Journal of Political Science* 59(4): 866-879.

Diermeier, D., C. Prato, and R. Vlaicu (2020) "Self-Enforcing Partisan Procedures," *Journal of Politics* 82(3): 937-954.

Duca, J.V. and J. Saving (2016) "Income Inequality and Political Polarization: Time Series Evidence Over Nine Decades," *Review of Income and Wealth* 62(3): 445-466.

Feddersen, T. and F. Gul (2015) "Polarization and Income Inequality: A Dynamic Model of Unequal Democracy," working paper.

Fiorina, M.P. and S.J. Abrams (2008) "Political Polarization in the American Public," *Annual Reviews of Political Science* 11: 563-588.

Garand, J. (2010) "Income Inequality, Party Polarization, and Roll-Call Voting in the U.S. Senate," *Journal of Politics* 72(4): 1109-1128.

Gerber, A.S. and D.P. Green (2000) "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment," *American Political Science Review* 94(3): 653-663.

Großer, J. and T. Palfrey (2014) "Candidate Entry and Political Polarization: An Anti-median Voter Theorem," *American Journal of Political Science* 58(1): 127-143.

Gerring, J. (1998) *Party Ideologies in America, 1828-1996*, Cambridge, UK: Cambridge University Press.

Hetherington, M.J. (2001) "Resurgent Mass Partisanship: The Role of Elite Polarization," *American Political Science Review* 95(3): 619-631.

Herrera, H., D.K. Levine, and C. Martinelli (2008) "Policy Platforms, Campaign Spending, and Voter Participation," *Journal of Public Economics* 92(3–4): 501-513.

Jacobson, G.C. (2012) "The Electoral Origins of Polarized Politics: Evidence from the 2010 CCES," *American Behavioral Scientist* 56(12): 1612-1630.

Kamada, Y. and F. Kojima (2014) "Voter Preferences, Polarization, and Electoral Politics," *American Economic Journal: Microeconomics* 6(4): 203-236.

Krasa, S. and M. Polborn (2010) "Competition between Specialized Candidates," *American Political Science Review* 104(4): 745-765.

Krasa, S. and M. Polborn (2014a) "Policy Divergence and Voter Polarization in a Structural Model of Elections," *Journal of Law and Economics* 57(1): 31 - 76.

Krasa, S. and M. Polborn (2014b) "Political Competition in Legislative Elections," working paper.

Levendusky, M. (2009) *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans.* Chicago: University of Chicago Press.

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

Nokken, T.P. and Poole, K.T. (2004) "Congressional Party Defection in American History," *Legislative Studies Quarterly* 29: 545–568.

Palfrey, T. (1984) "Spatial Equilibrium with Entry," *Review of Economic Studies* 51(1): 139-56.

Patty, J.W. (2008) "Equilibrium Party Government," *American Journal of Political Science* 52(3): 636-655.

Patty, J.W. and E.M. Penn (2017) "Uncertainty, Polarization, and Proposal Incentives Under Quadratic Voting," *Public Choice* 172: 109-124.

Piketty, T. and E. Saez (2003) "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics* 118(1): 1-39.

Poole, K.T. and H. Rosenthal (1984) "The Polarization of American Politics," *Journal of Politics* 46(4): 1061-1079.

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

Shor, B. and N. McCarty (2011) "The Ideological Mapping of American Legislatures," *American Political Science Review* 105(3): 530-551.

Snyder, J.M. and M.M. Ting (2002) "An Informational Rationale for Political Parties," *American Journal of Political Science* 46(1): 90-110.

Spenkuch J.L. and D. Toniatti (2016) "Political Advertising and Election Outcomes," working paper.

Van Weelden, R. (2013) "Candidates, Credibility, and Re-election Incentives," *Review of Economic Studies* 80(4): 1622-1651.

Van Weelden, R. (2015) "The Welfare Implications of Electoral Polarization," *Social Choice and Welfare* 45(4): 653-686.

Wittman, D. (1983) "Candidate Motivation: A Synthesis of Alternative Theories," *Amer-*

ican Political Science Review 77(1): 142-157.