

Policy divergence and voter polarization in a structural model of elections*

Stefan Krasa[†]

Mattias Polborn[‡]

October 29, 2012

Abstract

We analyze the relation between voters' issue preferences, the candidates' policy positions and voter behavior, using a simple spatial model of voter preferences. The size of policy differences between the candidates affects how important the voters' preferred policy positions are for their vote choices. Conversely, how vote choices depend on issue preferences provides information about the direction and extent of policy divergence. Based on the model, we propose a new method that simultaneously estimates the distribution of voter preferences and the extent of policy divergence between the candidates' platforms. Using data from multiple elections, we can separate changes in the electorate's polarization into those driven by *voter radicalization* and those due to increased *sorting*, stemming from policy divergence.

Keywords: Polarization, differentiated candidates, policy divergence, ideology, voter migration.

JEL Code: D72.

*We would like to thank seminar participants at the University of Rochester (Wallis Conference 2011), Princeton University, University of Cologne and University of Bielefeld for helpful comments.

[†]Department of Economics, University of Illinois, 1407 W. Gregory Dr., Urbana, IL, 61801. E-mail: skrasa@uiuc.edu

[‡]Department of Economics and Department of Political Science, University of Illinois, 1407 W. Gregory Dr., Urbana, IL, 61801. E-mail: polborn@uiuc.edu.

1 Introduction

Policy convergence or divergence is arguably the central question in political economy models of elections. The seminal contribution of Downs (1957) predicts equilibrium platform convergence, while there is a large number of variations of the spatial model of electoral competition that develop different reasons for policy divergence.¹ To better evaluate these theories of what drives party platform choices in electoral competition requires a method of measuring party divergence. While it is always difficult to assess whether policy platforms in reality are “surprisingly similar” or “surprisingly dissimilar” in any absolute sense, a question that is better defined is whether policy divergence increased or decreased over time.

There are two standard ways of measuring candidates’ political positions. The DW-NOMINATE method, based on Poole and Rosenthal (1984, 1985, 2000), allows to measure the relative positions of members of a legislature by comparing their respective votes on many different bills. By construction, this method cannot provide results for candidates that do not vote on the same set of issues (e.g., for the candidates for executive office such as the Presidency unless both candidates served in the same legislative body before running).

Second, one can directly ask voters about their assessment of the candidates’ positions. For example, the American National Election Survey (ANES) asks respondents, among many other questions, to classify the major candidates on a seven point scale that goes from “extremely liberal” (1) to “extremely conservative” (7). A problem with this approach is that, while the spatial “left-right” framework is second nature for political economists and many political scientists, there are many ordinary voters who appear uneasy to use the abstract framework of a spatial model to place candidates. For example, 23% percent of respondents placed Obama strictly to the right of McCain in 2008. For this reason, a measure that relies on concrete policy questions to measure voter preferences and candidate positions appears preferable.

In this paper, we propose a new method of estimating candidate positions that is based on a simple spatial model of voter preferences. Furthermore, we show how candidate position choices affect how “polarized” voters appear with respect to their political issue preferences.

In our model, voters have preferred policy preferences on a number of economic and non-economic (“cultural”) issues. In addition, each voter also receives a non-policy payoff from each candidate, which

¹This literature is too large to cite exhaustively. Assumptions that may generate policy divergence include policy motivation (e.g., Wittman 1983, Calvert 1985, Martinelli 2001, Gul and Pesendorfer 2009); entry deterrence (e.g., Palfrey (1984), Callander (2005)); incomplete information among voters or candidates (e.g. Castanheira 2003, Bernhardt, Duggan, and Squintani 2006, Callander 2008); and candidates with differentiated abilities (e.g., Soubeyran 2009, Krasa and Polborn 2010).

captures, like in a probabilistic voting model (e.g., Hinich 1978; Lindbeck and Weibull 1987; Dixit and Londregan 1995; Banks and Duggan 2005), both a systematic component (e.g., the candidate's competence) and an idiosyncratic component (e.g., how likable a voter finds each candidate).

The size of policy differences between the candidates affects how important the voters' preferred policy positions are for their vote choices. To see this, suppose that initially both parties take the same position (or very close ones). In this case, voters receive the same policy utility from both parties, and therefore only their idiosyncratic likes and dislikes decide how they vote. Now suppose that the two parties' positions move farther apart: Say, Democrats become pro-choice, while Republicans become pro-life. In this case, a voter's preferred position on the abortion issue should matter much more, in the sense of being a better predictor of vote choice, than in the initial situation. Note that the change in voter behavior here depends on changes in the platforms of the two parties, rather than on changes in the distribution of voter ideal points. More generally, our framework can be used to analyze how the parties' policy proposals and the distribution of voters' policy preferences affect voting behavior.²

The extent of policy divergence on different issues also determines along which lines the electorate appears divided. In a situation in which the two parties have very similar economic positions while differing substantially on "cultural" issues, voters primarily divide according to their cultural preferences: A large percentage of social liberals votes for Democrats, and a large percentage of social conservatives votes for Republicans. If, in another situation, the two parties have very similar economic positions while differing substantially on "cultural" issues, voters will primarily divide according to their cultural preferences: A large percentage of economic liberals will vote for Democrats, and a large percentage of economic conservatives will vote for Republicans. The fact that voting behavior seems to depend more on social preferences in the first scenario does not mean that voters "care more about social issues" or "are more extreme on social issues" in the first scenario (compared to the second one). In fact, it is perfectly possible that the distribution of voter issue preferences is identical in both scenarios. If the distribution of voter preferences changes, e.g. probability mass shifting from the middle of the distribution to more extreme positions, this is a completely separate effect that leaves the estimated behavior of any given ideological type unchanged, though it does affect how many voters "cross-over" to the other party.

While we do not observe party positions directly, we can see how voters' ideal positions stochastically

²In the model, we remain agnostic as to what determines the parties' policy positions or even whether candidates play a pure or mixed strategy equilibrium because this is irrelevant for our main question.

influence their voting behavior, i.e. whether they vote for Republicans or Democrats. Using this information in our model allows us to recover the distance between the parties' positions on different issues (up to a normalization constant). When we apply the model to U.S. Presidential elections from 1972 to 2008, we obtain clear evidence of increasing policy divergence between parties, primarily on cultural issues, but also on economic issues. As a consequence, cultural preferences become, relatively to economic preferences, more important predictors of voting behavior.

Our second main contribution is to analyze the question of political "polarization" in our framework. While many commentators diagnose a sharp and increasing ideological divide that splits the U.S. electorate,³ both the popular press and the existing literature are somewhat unclear about what exactly constitutes polarization and how one can measure this concept. We think of (ideological) polarization as a measure of how intensely voters' care on average about whether the Democrat or the Republican candidate is elected, and operationalize this notion as the degree to which voters' candidate choices depend on their preferred issue positions. For example, we would think of a situation in which the two candidates choose the same platform and in which voters' candidate choices consequently do not depend on their issue preferences, as not ideologically polarized. In contrast, a situation with strong policy divergence leads to intense preferences among most voters, and their vote choices will depend strongly on their issue preferences; this society appears ideologically polarized.

In general, this observation implies that policy divergence and the apparent polarization in a society are intimately connected, and this is the reason why we need to analyze both problems in the same framework. Looking at the intertemporal change of our polarization measure, we can determine whether U.S. voters are indeed more politically polarized today than a generation ago, and if so, is party platform divergence, a change in the voters' preferences, or both are responsible for this. Because our estimation procedure provides a distribution of voters' ideal points and the positions of candidates, in different elections, we can synthetically separate and quantitatively estimate the importance of the two potential reasons for changes in the overall polarization measure. In a first thought-experiment, we fix the candidates at their positions in a previous election, and look at only those changes that arise from changes in the distribution of voter ideal points alone. We call this effect "voter radicalization." Second, we fix the electorate of an earlier election year and see how this constant set of voters reacts to the observed change in the parties' positions. We call this effect "sorting." The data show a substantial increase in overall polarization between the 1970s

³For example, the Economist ("On His High Horse," November 9, 2002 issue) writes that "the 50-50 nation appears to be made up of two big, separate voting blocks, with only a small number of swing voters in the middle."

and today. Most of this increase is due to sorting, but there is also evidence for a substantial increase in radicalization, especially in the last decade.

In the next section, we set out our model. In Section 3, we define our key concepts, show how they correspond to the model and provide the theoretical basis for the estimation. In Sections 4, 5 and 6, we apply our methods to National Election Survey data from U.S. Presidential elections between 1972 and 2008, and discuss the implications of our results for theoretical models of policy divergence. Section 7 concludes. The Appendix contains proofs and a generalized model.

2 Model

Two candidates, labeled D and R , are endowed with a cultural-ideological position $\delta_P \in [0, 1]$, $P \in \{D, R\}$, an economic position g_P that denotes the quantity of a public good that the candidate provides if elected, and an associated cost of public good provision c_P .

Each voter is characterized by his cultural ideology $\delta \in [0, 1]$; a parameter $\theta \in [0, 1]$ that determines his preferences for public goods, and a parameter $\xi_P \in \mathbb{R}$ that measures the impact of the personal charisma of the candidate $P = D, R$ on the voter. Specifically, a voter's utility from candidate P is given by

$$u(\delta, \theta, \xi_P) = \theta v(g_P) - c_P - (\delta - \delta_P)^2 + \xi_P. \quad (1)$$

Note that $v(\cdot)$ is an increasing and strictly concave function that is the same for all voters. Since a voter's gross utility from public goods is $\theta v(g)$, high θ -types receive a higher payoff from public goods and thus, their preferred public good provision level, accounting for the cost of provision, is higher than for low θ -types.⁴ We assume that there is a continuous distribution of $(\delta, \theta, \xi_D, \xi_R)$ in the electorate, that $\theta \in [0, 1]$,⁵ and that $\xi \equiv \xi_R - \xi_D$ is independent of θ and δ .

For simplicity of exposition, we have presented the model with one economic and one cultural dimension. In the Appendix, we describe how the model can be modified for an arbitrary number of ideological issues. Also, our focus in this paper is on analyzing the consequences of policy divergence for voter behav-

⁴We could generalize the utility function to $u(P, g) = \theta v(g) - c_P - s(\delta - \delta_P)^2 + \xi_P$, where $s > 0$. The case $s = 1$ corresponds to (1), and higher s means that voters put more emphasis on cultural issues. By setting $\chi = \sqrt{s}(\delta - \bar{\delta}) + \bar{\delta}$, for arbitrary $\bar{\delta}$ we can write the new utility function as $u(P, g) = \theta v(g) - c_P - (\chi - \chi_P)^2 + \xi_P$, which is exactly the same form (1) (just with χ replacing δ). Thus, our assumption that the parameter multiplying the ideological loss $(\delta - \delta_P)^2$ is one is without loss of generality.

⁵This is just a normalization because $v(\cdot)$ can take arbitrary values.

ior. Thus, we remain agnostic as to which model describes the candidate’s policy choices; we simply take them as exogenously given.⁶ For example, Krasa and Polborn (2011) analyze endogenous policy choice in the same framework. However, from the perspective of the present paper, all that matters is that voters observe the positions of the two candidates and vote for the candidate who provides them with a higher utility. Whether candidates are exogenously committed to particular positions from the outset, or can choose which policies to commit to before the election, is irrelevant.

3 Analysis of the Model

3.1 The Cutoff Line

A voter is indifferent between the two candidates if and only if $\theta v(g_D) - c_D - (\delta - \delta_D)^2 + \xi_D = \theta v(g_R) - c_R - (\delta - \delta_R)^2 + \xi_R$, which implies

$$-2\delta(\delta_R - \delta_D) + (v(g_D) - v(g_R))\theta = c_D - c_R - (\delta_R^2 - \delta_D^2) + \xi. \quad (2)$$

We assume that the Democrat provides more of the public good for a higher tax cost (i.e., $g_D \geq g_R$ and $c_D \geq c_R$), and that the Republican is to the right of the Democrat on cultural issues (i.e., $\delta_R \geq \delta_D$).⁷

For any given value of ξ , if $g_D = g_R$, the line of indifferent or *cutoff voters* in a (δ, θ) -space is vertical. Intuitively, if Democrat and Republican provide the same amount of public goods, then only the voters’ ideological preferences (δ) matter for their vote choice, while the voters’ economic preference (θ) is immaterial. If, instead, $g_D > g_R$, the cutoff value for θ is given by

$$\theta(\delta, \xi, g_D, g_R) = \frac{2\delta(\delta_R - \delta_D) + c_D - c_R - (\delta_R^2 - \delta_D^2) + \xi}{v(g_D) - v(g_R)}. \quad (3)$$

Equation (3) is a straight line in the δ - θ space, and has a positive slope. Intuitively, if the Democrat provides more public goods than the Republican, then a voter is indifferent between the candidates either if he is socially liberal, but wants lower spending on public goods (i.e., low δ and low θ), or if he is socially conser-

⁶Note that this approach does not generate an endogeneity problem in the empirical analysis, because at the time the voters make their decisions, the candidates have chosen their positions.

⁷From a theoretical point of view, these are mere normalizations: We can simply call the candidate who provides more public good the “Democrat,” and measure δ in a way that the Democrat’s position is weakly to the left of the Republican’s. These normalizations make sense in the U.S. context.

vative, but likes substantial government spending on public goods (i.e., high δ and high θ). Higher types of θ are more likely to vote for the Democrat, and for any given economic preference type θ , higher δ -types are more likely to vote for the Republican.

3.2 Determining voter types

Our next objective is to translate a respondent's answers to the survey questions into a position in the δ - θ -space, and a probability of voting Republican. The separating line (3) is determined by the candidates' positions and may therefore change from one election to the next. In particular, the slope, k , and the intercept, a are given by

$$k = \frac{2(\delta_R - \delta_D)}{v(g_D) - v(g_R)}, \quad a = \frac{c_D(g_D) - c_R(g_R) - (\delta_R^2 - \delta_D^2) + \bar{\xi}}{v(g_D) - v(g_R)}. \quad (4)$$

where $\bar{\xi} = E[\xi]$. Define

$$\varepsilon = \frac{\xi - \bar{\xi}}{v(g_D) - v(g_R)} \quad (5)$$

We assume that ε is normally distributed with standard deviation σ (given the normalization in (5), the mean of ε is 0). Equations (3), (4) and (5) imply that a citizen votes Republican if and only if

$$\theta - k\delta - a - \varepsilon < 0. \quad (6)$$

Let $X_i, i = 1, \dots, n$ and $Y_i, i = 1, \dots, m$ be random variables that describe the answers to survey questions on cultural and economic issues, respectively. We assume that $\delta = \sum_{i=1}^n \lambda_i X_i$ and $\theta = \sum_{i=1}^m \mu_i Y_i$, where, of course, the λ_i and μ_i are parameters to be estimated.

We normalize X_i and Y_i such that (i) the lowest and highest realizations for each question are 0 and 1; (ii) high values on X_i and Y_i increase the estimated value of δ and θ , respectively (i.e., we code answers such that all λ_i and μ_i are non-negative).⁸ Finally, we normalize $\sum_{i=1}^n \lambda_i = 1$ and $\sum_{i=1}^m \mu_i = 1$ so that $\theta, \delta \in [0, 1]$, to keep the distribution of θ and δ comparable over time. This normalization is without loss of generality because multiplying all variables in (6) by a positive constant does not change whether (6) is satisfied.⁹

Let $\Phi(\cdot)$ denote the cdf of a normal distribution with mean 0 and standard deviation 1. Then (6) implies

⁸Clearly, this can be done by defining a new random variable $\hat{X}_i = 1 - X_i$ ($\hat{Y}_i = 1 - Y_i$) if λ_i (or μ_i) is negative.

⁹In the estimation, multiplying all variables in (6) by the same constant leaves the parameter estimate for k unchanged and multiplies the estimate of the standard deviation of ε accordingly.

that the probability that a voter votes Republican is given by

$$\Phi\left(\frac{1}{\sigma}\left[k\sum_{i=1}^n\lambda_iX_i - \sum_{i=1}^m\mu_iY_i + a\right]\right). \quad (7)$$

We now describe how the model can be used to identify changes in the distance between the candidates' platforms. Taking the standard deviation on both sides of (5) we get

$$\sigma = \frac{\sigma_\xi}{v(g_D) - v(g_R)} \quad (8)$$

where σ_ξ is the standard deviation of ξ . We assume that σ_ξ does not change over time, but make no assumption about the average value of ξ in the population, i.e. the average net valence of candidates is allowed to vary over time.¹⁰

Using (4) implies

$$\delta_D - \delta_R = \frac{\sigma_\xi k}{2\sigma}, \text{ and } v(g_D) - v(g_R) = \frac{\sigma_\xi}{\sigma} \quad (9)$$

We can use equations (17) and (18) in Theorem 1 (in Section 3.4 below) to estimate the values σ and k for different years. This allows us to identify both the cultural and economic difference in the candidates' platforms, if we normalize the policy difference $v(g_D) - v(g_R)$ in a base year.

3.3 Polarization, Radicalization and Sorting

“Polarization” is a central issue in the analysis of American political behavior. Many commentators diagnose a sharp and increasing partisan divide that splits the U.S. electorate. For example, the Economist writes that “the 50-50 nation appears to be made up of two big, separate voting blocks, with only a small number of swing voters in the middle”, and that “America is more bitterly divided than it has been for a generation”.¹¹

While many pundits and political scientists write about “polarization”, there is no general agreement on a formal definition of this concept. Intuitively, it does not make sense to define polarization by how close

¹⁰In a model that analyzes data from only one year, the assumption that the residual error is drawn from a standard normal distribution is a mere normalization because the objective function (7) is homogeneous of degree zero in σ and the regression parameters, and thus σ can be normalized without loss of generality. In a multi-period model, the model identifies changes in coefficients only *relative to the distribution of the error term*. Assuming that σ_ξ is constant over time allows us to skip the part in italics when interpreting the change of regression coefficients (or functions of regression coefficients) over time. This is a standard assumption when the analysis is based on a comparison of regression coefficients over time (e.g. Bartels 2006, McCarty, Poole, and Rosenthal 2006) and usually not even discussed.

¹¹“On His High Horse,” November 9, 2002 issue and “America’s Angry Election,” January 3, 2004 issue, respectively.

the election outcome is to a “50-50” split – that feature is more appropriately defined as competitiveness or closeness. Not every close election is meaningfully characterized as polarized; for example, consider the equilibrium of the original Downsian model in which both candidates choose the same position and where therefore all voters are indifferent between candidates. If, in the case of indifference, each voter flips a coin to decide which candidate to vote for, the election result in a large electorate is very close, but it clearly would not make sense to call this a polarizing election.

A meaningful notion of polarization requires a certain intensity of preference among many voters (the “bitterly divided” part of the quote above). A natural notion of political polarization from an economist’s point of view would be to measure each voter’s willingness to pay for a victory of their preferred candidate and aggregate the absolute values of this willingness to pay. Evidently, this concept cannot be operationalized directly.

However, the relative preference intensity is reflected in voter behavior. Consider a setting in which each voter is characterized by a multidimensional type vector (e.g., ideological positions, gender, race, ethnicity, idiosyncratic sympathy for a candidate); some parts of each voter’s type vector are observable. When people care so intensely that they appear “polarized” along a certain observable dimension in the type space, this part of their type is a very good predictor of their behavior (and vice versa).

For example, consider a hypothetical polity in which voters care about taxes, abortion and candidate personality. If candidates differ only on taxes, and not by much, then we would think of this society as not very polarized: While voters’ economic preferences are somewhat predictive of their vote choice, there are many swing voters who are willing to “cross over” if they prefer the personality of their ideologically less preferred, and we can conclude that voters’ political preference intensities are small. Moreover, the small extent of political polarization in this society is only along the economic issue. In contrast, if candidate positions diverge, voters’ average net preference for their respective preferred candidate increases.

To formalize this concept, suppose that we have to predict the voting behavior of a large group of voters in a tight election. If we did not have any information about these voters, we could not do better than flipping a coin, and this would give us a 50 percent “success quota.” Using information about a voter’s ideology enables us to make better predictions. If a voter’s ideology is below (above) the separating line and we predict him to vote Republican (Democrat), then the probability that the prediction is correct is $\Phi\left(\frac{1}{\sigma_t}[k_t\delta_i - \theta_i + a_t]\right)$, where (k_t, a_t, σ_t) denote the parameters for a separating line for year t . When we average this measure over all voters, we have a measure of how important political issue preferences are for

predicting voting behavior.

Note that a problem could arise in lopsided elections. For example suppose that 70 percent of voters vote for the Republican candidate in an election because that candidate has a large expected valence $\bar{\xi}$. Then even a completely uninformed guesser could achieve a 70 percent success quota (by guessing that each voter votes Republican). To avoid this problem, we adjust the valence such that the election would have ended in a tie. More formally, we find a new intercept a'_t such that the weighted vote share of the Democrat (and Republican) is exactly 1/2, i.e. $(1/I) \sum_i \Phi\left(\frac{1}{\sigma_t}[k_t \delta_i - \theta_i + a'_t]\right) = 0.5$. We then measure the quality of information about political positions by how much the success quota of our forecasting system lies above the success quota of a pure coin flip:

$$\Psi_t = \frac{2}{I} \sum_{i=1}^I \left| \Phi\left(\frac{1}{\sigma_t}[k_t \delta_i - \theta_i + a'_t]\right) - 0.5 \right|. \quad (10)$$

Note that $\left| \Phi\left(\frac{1}{\sigma_t}[k_t \delta_i - \theta_i + a'_t]\right) - 0.5 \right|$ is the increase in the success probability relative to a pure coin flip, and the factor 2 in front normalizes Ψ such that it lies between 0 and 1. For example, if knowledge of political preferences allows to correctly forecast 80 percent of voters, then this is $2(0.8 - 0.5) = 60\%$ better than a pure coin flip.

If $\Psi = 1$, society is extremely divided along ideological lines: Every conservative votes Republican, and every liberal votes Democratic. This means that most voters would know which party they will vote for before they know who are the actual candidates of each party – they are not going to give the other party’s candidate a chance to convince them to switch parties in this election, and there are no “swing voters.” In contrast, if $\Psi = 0$, knowledge of a voter’s issue preferences does not help to predict voting behavior – all voters are ex-ante open to both candidates.

Changes in Ψ over time may arise for two distinct reasons. First, candidates’ platforms may be more distinct, generating stronger preference intensities among voters. Second, voters themselves may become more extreme in their political views (i.e., their ideal points).

Figure 1 illustrates these two effects. In the left panel, the distribution of voter ideal points remains constant, but the “isoprobability lines” — the lines along which the probability of voting for a candidate is constant — move closer to the 50% line which occurs because of policy divergence. Simple geometry shows that the distance from the 50% line to any other “isoprobability” line, such as the 75% line in the

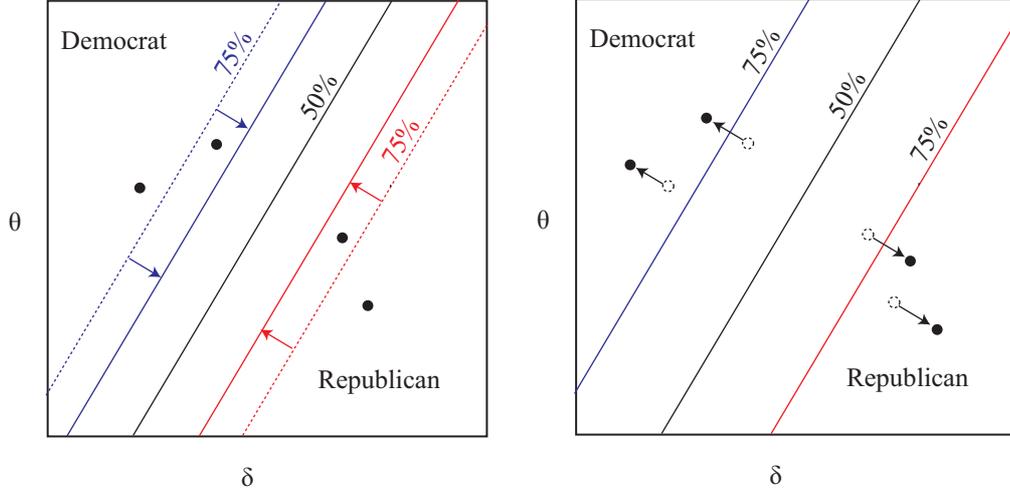


Figure 1: Increasing Polarization through Sorting and Radicalization

graph, is proportional to $\sigma / \sqrt{1 + k^2}$. Thus, (4) and (8) imply that the distance is proportional to

$$\frac{\sigma_{\xi}}{\sqrt{[v(g_D) - v(g_R)]^2 + 4(\delta_R - \delta_D)^2}}.$$

As a consequence, an increase in the distance between the candidates' policy positions moves the isoprobability lines closer together in the left panel of Figure 1, which results in an increase of Ψ . We refer to this effect as *sorting*. Voters ideological positions are unchanged but their voting behavior is more predictable since the candidates offer more distinct policy platforms.

The right panel of Figure 1 illustrates the second reason why polarization may increase: Voters policy positions become more extreme, so that it is easier to predict how people vote. We refer to this effect due to the movement of voter ideal points as *radicalization*.

To formally separate sorting from radicalization, let $\Psi(t, t')$ denote the polarization for the electorate of year t if the politicians' positions are as in year t' . The total change in polarization in year t from the previous election in year $t - 4$ is $\Delta\Psi_t = \Psi(t, t) - \Psi(t - 4, t - 4)$. We call $\Delta S(t) \equiv \Psi(t - 4, t) - \Psi(t - 4, t - 4)$ the level of sorting in year t , taking as given the base electorate of the last election. The remaining change in Ψ , given by $\Delta R(t) = \Psi(t, t) - \Psi(t - 4, t)$, captures the effect of radicalization due to the movement of voter ideal points.

It is interesting to note that changes in an average willingness to pay measure of polarization would also

be separable in two analogous parts: A given voter's willingness to pay for the election of his preferred candidate changes as the candidates' positions change; this effect is analogous to our sorting effect. Alternatively, an average willingness to pay measure of polarization could increase, holding fixed the candidates' positions, because voters radicalize and would be (on average) willing to pay more for the election of their favorite candidate; this effect is analogous to our radicalization effect.

Finally, note that we can apply the concepts of polarization, sorting and radicalization to the full set of issues (which we will do in Section 5.3), or only to a subset of issues. For example, the latter approach would allow us to make statements such as "the U.S. electorate has become more polarized with respect to economic issues," or whatever other subset of issues we choose to analyze.

3.4 Estimation Procedure

In order to get the best estimate of voters' values of δ and θ , we estimate λ and μ using pooled data from several elections. Because candidate platforms change from one election to the next, this means that we must allow that k and σ change over time and thus index them by the year of the election. Let D_t , $t = 1, \dots, s$ be the year dummy for year $t = 1, \dots, s$ (i.e., $D_t = 1$ if the observation occurred in year t , and 0 otherwise). Then (7) generalizes to

$$\Phi \left(\sum_{t=1}^s \frac{D_t}{\sigma_t} \left[\left(\sum_{i=1}^s D_t k_t \right) \left(\sum_{i=1}^n \lambda_i X_i \right) - \sum_{i=1}^m \mu_i Y_i + \sum_{t=1}^s D_t a_t \right] \right). \quad (11)$$

In order to determine k_t , a_t , σ_t , $t = 1, \dots, s$, λ_i , $i = 1, \dots, n$, and μ_i , $i = 1, \dots, m$, we first estimate the model in which the probability of voting Republican is given by

$$\Phi \left(\left(1 + \sum_{t=2}^s \alpha_t D_t \right) \left(\sum_{i=1}^n \tilde{\lambda}_i \tilde{X}_i \right) - \left(1 + \sum_{t=2}^s \rho_t D_t \right) \left(\sum_{i=1}^m \tilde{\mu}_i \tilde{Y}_i \right) + \sum_{t=1}^s \tilde{a}_t D_t \right), \quad (12)$$

where there are no restrictions on the $\tilde{\lambda}_i$, and $\tilde{\mu}_i$, i.e., they could be negative or greater than 1. \tilde{X}_i and \tilde{Y}_i are the responses to the survey questions, solely normalized to be between 0 and 1, but absent the additional requirement that higher realizations of the response to each question increase δ and θ .

Denote by $d_{t,\ell}$, $\tilde{x}_{i,\ell}$, and $\tilde{y}_{i,\ell}$ observation ℓ of random variables D_t , \tilde{X}_i , and \tilde{Y}_i , respectively. Let

$$z_\ell = \left(1 + \sum_{t=2}^s \alpha_t d_{t,\ell} \right) \left(\sum_{i=1}^n \tilde{\lambda}_i \tilde{x}_{i,\ell} \right) - \left(1 + \sum_{t=2}^s \rho_t d_{t,\ell} \right) \left(\sum_{i=1}^m \tilde{\mu}_i \tilde{y}_{i,\ell} \right) + \sum_{t=1}^s \tilde{a}_t d_{t,\ell}, \quad (13)$$

and let $v_\ell = 1$ if the voter in observation ℓ votes Republican, and $v_\ell = 0$ if he votes Democrat. To estimate $\alpha_i, \beta_i, \tilde{\lambda}_i, \tilde{\mu}_i$, and \tilde{a}_i , we maximize the log-likelihood function, i.e., solve

$$\max_{\{\alpha_i, \rho_i | i=2, \dots, s\}, \{\tilde{a}_i | i=1, \dots, s\}, \{\tilde{\lambda}_i | i=1, \dots, n\}, \{\tilde{\mu}_i | i=1, \dots, m\}} \sum_{\ell=1}^L v_\ell \ln \Phi(z_\ell) + (1 - v_\ell) \ln (1 - \Phi(z_\ell)). \quad (14)$$

We use Newton's method to determine a zero of the first order condition of this maximization problem. Note that, in contrast to a standard probit model, z_j is not a linear function of the model parameters. This generates some numerical challenges, as the region of convergence is relatively small, thus requiring a good start value.¹² The computer code for performing the estimation can be obtained from the authors. Theorem 1 shows how the parameter estimates of (14) translate into parameters of the original model.

Theorem 1 *Define $\rho_1 = \alpha_1 = 1$. Let α_t, ρ_t and \tilde{a}_t for $t \in \{1, \dots, s\}$; $\tilde{\lambda}_i, i \in \{1, \dots, n\}$; $\tilde{\mu}_i, i \in \{1, \dots, m\}$, be the parameters of the modified model in (12). Then the parameters of the original model (11) are determined as follows:*

1. δ and θ are given by

$$\delta = \frac{\sum_{i=1}^m [\tilde{\lambda}_i \tilde{X}_i - \min\{\tilde{\lambda}_i, 0\}]}{\sum_{i=1}^m |\tilde{\lambda}_i|}, \quad \theta = \frac{\sum_{i=1}^n [\tilde{\mu}_i \tilde{Y}_i - \min\{\tilde{\mu}_i, 0\}]}{\sum_{i=1}^n |\tilde{\mu}_i|}. \quad (15)$$

2. The weights of cultural and economic issues are given by

$$\lambda_i = \frac{|\tilde{\lambda}_i|}{\sum_{i=1}^n |\tilde{\lambda}_i|}, \quad \mu_i = \frac{|\tilde{\mu}_i|}{\sum_{i=1}^m |\tilde{\mu}_i|} \quad (16)$$

3. The standard deviation of the individual preference shock ε_t in period t is given by

$$\sigma_t = \frac{1}{(1 + \rho_t) \sum_{i=1}^m |\tilde{\mu}_i|} \quad (17)$$

4. The slope of the separating line in the (δ, θ) space in period t is

$$k_t = \frac{(1 + \alpha_t) \sum_{i=1}^n |\tilde{\lambda}_i|}{(1 + \rho_t) \sum_{i=1}^m |\tilde{\mu}_i|} \quad (18)$$

¹²We obtain such a start value by first optimizing over $\tilde{\lambda}_i, \tilde{\mu}_i$ and \tilde{a}_i , use the resulting solution as a start value for optimizing over α_i, ρ_i , and \tilde{a}_i . Starting from this value, convergence can be obtained for the complete optimization problem.

5. The vertical intercept of the separating line in the (δ, θ) space in period t is

$$a_t = \frac{\tilde{a}_t - (1 + \rho_t) \sum_{i=1}^m \min\{\tilde{\mu}_i, 0\} + (1 + \alpha_t) \sum_{i=1}^n \min\{\tilde{\lambda}_i, 0\}}{(1 + \rho_t) \sum_{i=1}^m |\tilde{\mu}_i|}. \quad (19)$$

After determining weights (λ, μ) for a set of base years, we can determine δ and θ by using (15). Different base years give slightly different results because which preferred policy positions are economically and culturally “conservative” (i.e., leaning towards the Republican position) may change over time. Different approaches have different advantages. Pooling all years gives us the largest data set and compares all years against a common benchmark. In contrast, focusing on late base years has the advantage of measuring people’s preferences in a way that is more consistent with what is considered economically and culturally liberal or conservative *today* as opposed to an average over the last generation, and this the approach that we will choose for the main part of the paper. However, it should be noted that our main results are not sensitive with respect to the choice of the base period.

For given values of (δ, θ) (i.e., obtained for (λ, μ) from fixed base years), estimating k , a , and σ for a single election year is much easier. In particular, the probability of voting Republican is given by

$$\Phi\left(\frac{1}{\sigma} [k\delta_i - \theta_i + a]\right). \quad (20)$$

We can estimate this model by first estimating

$$\Phi(\beta_\delta \delta_i - \beta_\theta \theta_i + \beta_a), \quad (21)$$

which is a standard probit model, and by then using the identities

$$\sigma = \frac{1}{\beta_\theta}, \quad k = \frac{\beta_\delta}{\beta_\theta}, \quad a = \frac{\beta_a}{\beta_\theta}. \quad (22)$$

Of course, if we compare the values that we get from a direct estimation of (14), and the values obtained by this simplified method in a year that is one of the base years used in the estimation of (14) (i.e., using the values of (δ, θ) obtained from the direct estimation), then we get identical values for k , σ and a .

4 Concepts and Data

We apply our model to U.S. Presidential elections from 1972 to 2008, using data from the American National Election Survey (henceforth NES). The advantage of the NES relative to (often media-financed) polls is that there is considerably more continuity in terms of the policy questions asked. We use all questions that were continuously available between 1972 and 2008 and indicate a voter's cultural or economic preferences.¹³

We group these questions into two policy areas, "economic" and "cultural" (i.e., everything else). Our method allows for splitting the questions into more areas, but a two-dimensional policy space allows for a nice graphical presentation of voter ideal points and voting behavior, and an easier interpretation of the relative importance of cultural and economic positions for vote choice.

We use the following questions in order to determine the cultural ideology index δ of a voter: Questions VCF0837/38 about abortion; question VCF0834 about the role of women in society; Questions VCF0206 and VCF0830, about the respondent's feeling towards blacks and affirmative action; Question VCF0213 about the respondent's feeling towards the U.S. military; Question VCF0130 about church attendance, which we use as a dummy with 1 for respondents who go to church weekly or almost every week. For economic preferences, we use Question VCF0809 on the role of the government in the economy; and Questions VCF0209 and VCF0210 about the respondent's feeling towards unions and "big business", respectively;

Of course, most of these questions are not questions about one narrowly-defined concrete policy issue that is constant over time. In fact, this likely occurs in any long-term data set: few questions about a very specific policy issue will remain topical for decades. However, the questions measure basic convictions that are very likely to relate to positions on the concrete policy issues of the day.¹⁴ A voter who felt negatively about the U.S. military in the 1970s was probably in favor of withdrawing from the Vietnam war, and a voter who felt negatively about the U.S. military in the last decade was probably in favor of withdrawing from the Iraq war. The concrete policy issues change, but the questions remain useful to measure basic convictions. Weekly church attendance may measure preferences on school prayer, subsidies for faith-based initiatives and other separation of church and state issues. The attitude towards unions and big business should be a good proxy for right-to-work legislation or business regulation in general.

¹³Because we need continuously available questions, we start our analysis in 1972, as moving to the 1960s would have meant losing a substantial number of questions. Conversely, moving the start date into the late 1980s and 1990s would expand the number of questions for which data are available, but at the cost of shortening the time series substantially.

¹⁴Also, voters will likely not base their candidate choice only on the candidates' positions about very specific policy issues, but rather on what they perceive to be the candidates' core convictions that will guide their respective decisions if elected.

Data on respondent’s demographic characteristics (such as gender and race) is available, but we have chosen not to use these variables as “policy positions.” A voter’s demographic characteristics may certainly influence his preferred positions; for example, women have on average a more liberal position on abortion rights than men, so if one did not know a voter’s preferences on abortion, including information on the voter’s gender is a useful proxy for preferred positions. However, since the NES has information on policy preferences, we prefer to use this information directly.¹⁵ Also, using demographic characteristics would make it harder to interpret our results. For example, suppose that we were to find that gender becomes a more important predictor of voting behavior. Since gender could plausibly correlate with both economic and non-economic policy preferences, this would not tell us anything definitive about the policy area in which the parties diverged.

We ignore the respondents’ partisan affiliation and self-placement on a liberal-to-conservative scale, both because of the problems with these measures raised in the introduction and because including such a measure would defy the purpose of our analysis. We want to know which policy-preferences (on both the economic and the cultural dimension) translate into a preference for the candidate of one of the parties. Regressing individuals’ vote choices for Democrats or Republicans on whether the individuals feel attached to either party, while done in many political science studies, is not very helpful for this objective. A similar comment applies to the respondents’ self-placement on a liberal-conservative scale.

5 Empirical Results

5.1 Finding the distribution of voter preferences (δ, θ)

The first step in the empirical analysis is to find the weights of different issue questions for the determination of the voters’ ideological positions. As described in detail in Section 3.4, we choose a set of base years and essentially pool the data from these years, and then take the relative magnitudes of the estimated regression coefficients as the weights. However, we have to take into account that there were different degrees of policy divergence in different elections, and the year dummies in (12) take care of this effect.

¹⁵Also, controlling for the respondent’s opinion about abortion and the role of women, the respondent’s gender does not provide much additional information about the voter’s preferences. In fact, we have run our regression including a number of demographic controls, and with some exceptions, they have turned out to be small and often insignificant. Also, dummies for the major religious groups (Protestants, Catholics) turned out to be very close to zero and statistically insignificant. This implies that, to the extent that Protestants are considerably more Republican than Catholics, the reason is that their average policy preferences differ – there is no direct effect of Catholicism.

It may be useful to remark that, if we were to choose just one year as the base period, then (12) specifies a standard probit model (since all data are derived from voters facing the same choice problem between the same pair of candidates). By pooling data from several elections, we base the calculation of these weights on more data which provides for some smoothing. However, pooling data from too many elections also has a drawback: It bases the notion of what positions are most important for the classification as an economic or social conservative on the behavior of voters many years ago, and what made a person economically or culturally conservative in the 1970s may be different from what we would think today. As a compromise, we choose the five elections between 1992 and 2008 as the base period that we use for the remainder of the analysis; however, we have checked that the qualitative results for policy divergence and polarization are robust to using other base periods such as 1972–2008 or 1972–1992.

Table 1 in the appendix reports the values and 95 percent confidence interval (obtained by using bootstrap resampling) of $\tilde{\lambda}$, $\tilde{\mu}$, λ , and μ for the two different pooled base periods, the first five elections (1972-1988) and the last five presidential elections (1992-2008). A person is more economically conservative (i.e., low θ) if he likes big business; dislikes unions; and does not feel that government should provide guaranteed jobs and a standard of living for everyone. The results for $\tilde{\lambda}$ and $\tilde{\mu}$ show the expected effects of political positions on voting behavior. A voter is more culturally conservative (i.e., high δ) if he likes the military; is against special government support for minorities; feels “less warm” towards blacks, believes that caring for the family is better for women than working outside the home; believes that abortion should be illegal; and attends church weekly or almost every week. With the exception of the role of women for the early base period, all coefficients are significant on the 95 percent level.

In terms of their weight for the determination of the economic index in the 1992 to 2008 period, the big business and government role question account for about one-quarter each, while the remaining half is determined by preferences on unions. Cultural preferences depend strongly on the respondent’s view of the military (about 30 percent weight; this is the only question on the whole nationalism/foreign policy complex that is consistently available throughout our observation period), the questions of race and affirmative action (about 40 percent) and the women-specific questions (about 25 percent). Note that weekly church attendance, while significant, has a surprisingly small weight, presumably because the opinions correlated with “Christian conservatism” are already reflected in the opinions expressed on the other issues.

Overall, the importance of different issues for the determination of the cultural and economic scores are relatively stable when comparing the earlier and the later period, though there are some interesting

changes: the importance of race issues decreased and that of women issues (role of women, abortion) increased.¹⁶ On the economic index, the importance of the union question increased relative to the big business question. Note that the real issue for robustness is not whether and to what extent the weights of different questions have changed, but whether the changed weights lead to qualitatively different results with respect to estimated party platform divergence, as well as radicalization and sorting. As we show in the online appendix, these results are qualitatively robust to the weights from different base periods.

5.2 Platform Differentiation

We now turn to the changes in the distance between the candidates' platforms. Recall from equation (9) in Section 3.2 that the model identifies changes in the policy distance relative to the distance in the base year. The base year is arbitrary, and we choose 1976 as base year since divergence on both policies is lowest in that year. Figure 2 displays the results for cultural and economic positions.

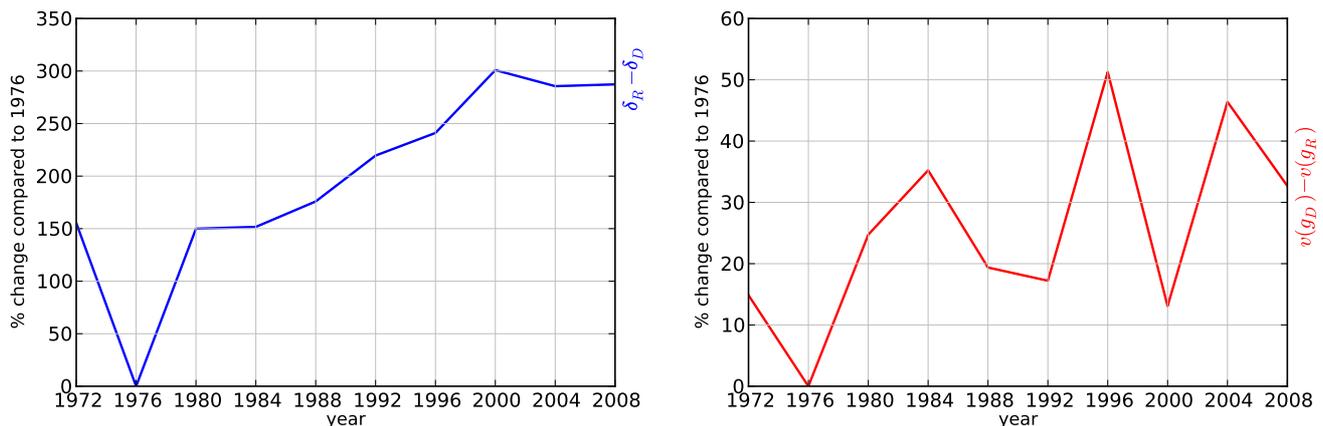


Figure 2: Cultural and economic policy divergence of candidates, 1972 to 2008

The difference between the two parties' cultural positions, $\delta_R - \delta_D$ increases by more than 200 percent in all years after 1992, and by about 300 percent in the last decade. For economic positions, the change in the distance between positions is considerably smaller; the maximum increase is about 50 percent in 1996. It should be noted, however, that our method only allows us to identify changes of the distance in

¹⁶The reader may, in particular, wonder why the importance of the seemingly quaint and today mostly uncontroversial "role-of-women" question has increased for the determination of social conservatism. However, it is exactly because an equal rights role of women is uncontroversial with most voters that a more conservative opinion on this issue has become a really strong signal for a respondent's cultural position.

cultural positions *relative to* the same distance in 1976, and many researchers have argued that the parties' positions on "moral issues" (a subset of our cultural issues here) were quite close to each other in the 1970s (e.g. Fiorina, Abrams, and Pope 2006; Ansolabehere, Rodden, and Snyder 2006), while the distance on economic issues may have been more substantial already in the base year.

We now turn to the question of how policy divergence affects voter behavior. Figure 3 provides some insights here. It displays the values of δ and θ for all voters, together with the voter's choice (red for Republican, blue for Democrat). The left panel is for the 1976 election, the right one for the 2004 election. In both panels, we have drawn the 50 percent separating line, i.e., voters on this line have an implied probability of voting Republican or Democrat that is exactly 1/2. Voters below and to the right of the separating line are more likely to vote for the Republican, while voters above and to the left of the line are more likely to vote for the Democrat.

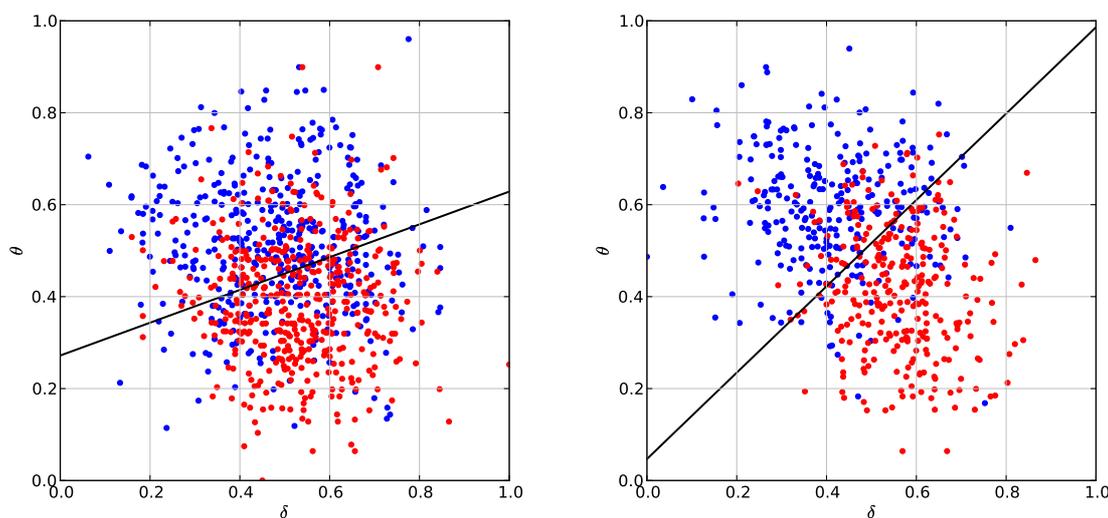


Figure 3: Voter preferences and vote choices in the 1976 (left) and 2004 (right) U.S. Presidential elections. Democratic voters in blue, Republican ones in red

Two features are evident from Figure 3. First, the ideological separation between Democrats and Republicans is much sharper in 2004 than in 1976. Clearly, this is what we would expect given that policy divergence, both on economic and on cultural issues, is substantially stronger in 2004. We will elaborate on this finding in Section 5.3.

Second, the slope of the dividing line, k , is low 1976: Voters split primarily along economic issues (with high θ types mostly voting for Carter, and low θ types mostly voting for Ford). In contrast, in 2004, the

separating line is considerably steeper and thus, to a higher degree along cultural lines, with social liberals primarily voting for Kerry, social conservatives for Bush. This is a consequence of the relatively stronger increase of policy divergence on cultural issues than on economic ones implies that the voters' cultural preferences δ become more important relative to their economic preferences θ .

We can interpret the slope k of the dividing line as a “marginal rate of substitution” between cultural and economic positions. That is, if an individual on the dividing line becomes one unit more culturally conservative, his economic liberalism needs to increase by k units in order for him to remain stochastically indifferent between the candidates. Remember that a change in k is *not* a consequence of voters directly putting more weight on cultural issues now than they did in the past, but rather follows from increased party divergence. Changes in the distribution of preferences are independent of these changes in k .

Figure 4 displays the development of the slope k for all years. After the initial decrease in k from 1972 to 1976, the relative importance of cultural issues starts to increase to reach a high point in 2000, and remaining relatively high afterwards. The confidence intervals in Figure 4 clearly indicate that, while election-to-election changes are often not statistically significant, the long-term trend definitely is statistically significant.

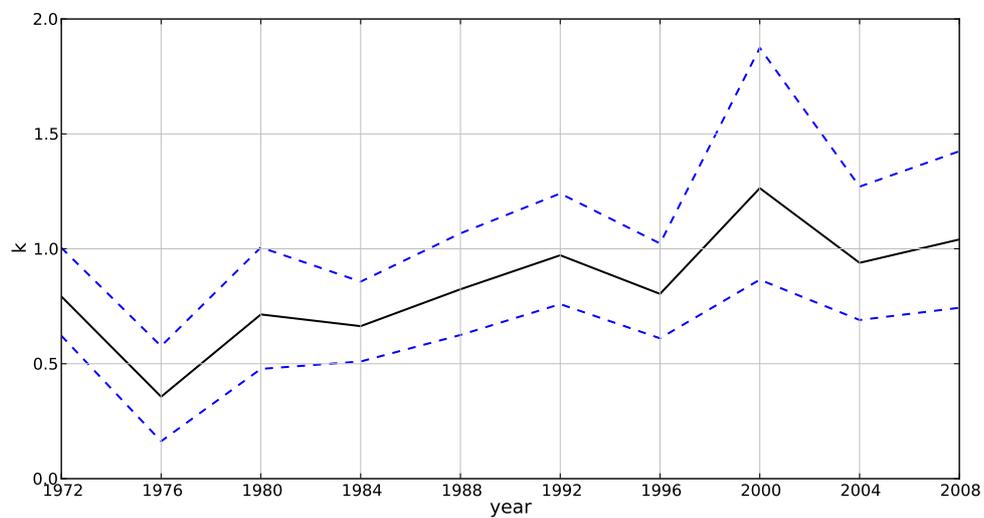


Figure 4: The development of k from 1972 to 2008, with 95% confidence intervals

5.3 Polarization , Radicalization and Sorting of the Electorate

We now return to the observation in the previous subsection that the increase in policy divergence implies that voters' policy preferences become a better predictor of their voting behavior. As proposed in Section 3.4, polarization Ψ is a useful formal measure of how well the voters in the ideology space are separated into voting blocks for Democrats and Republicans.

Figure 5 shows the development of Ψ over the last 10 presidential elections, and the parallels to cultural policy divergence in Figure 2 are quite obvious. Ψ decreases from 1972 to 1976 (to around 0.35), and then increases substantially throughout our observation period to end at a level of about 0.58. In other words, voters' political issue preferences are a substantially better predictor of their voting behavior in the 2000s than in the 1970s – knowing them allows about 65 percent better predictions in 2004 than it did in 1976.

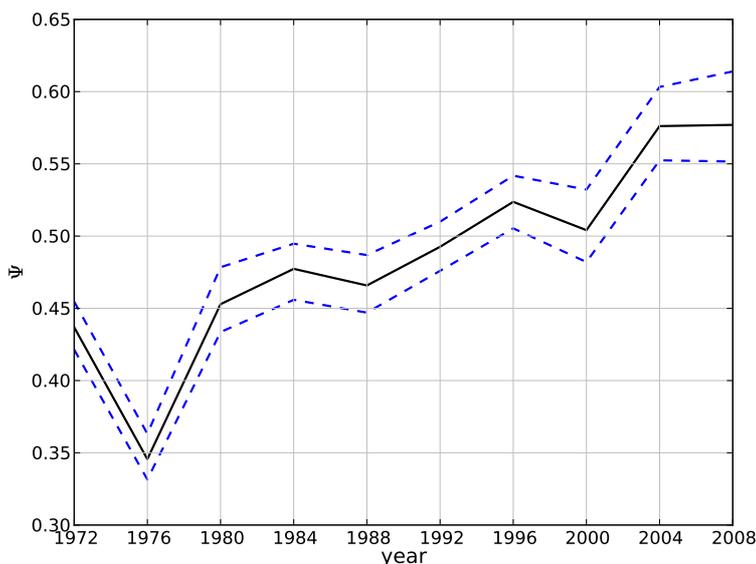


Figure 5: Polarization from 1972 to 2008, with 95% confidence intervals

As explained in Section 3.4, we can decompose the change in Ψ into the effect due to sorting and radicalization. Sorting isolates the effect of changes in platforms, holding fixed the distribution of political preferences in society at the level of the previous election in $t - 4$. Radicalization isolates the effect of a change in the voter preference distribution, holding fixed the platforms of the candidates.

Figure 6 plots $\Delta S(t)$ and $\Delta R(t)$ (defined in section 3.3 to decompose the change in polarization into radicalization and sorting). Since we do not have data for 1968 (i.e., the election before 1972), we start in

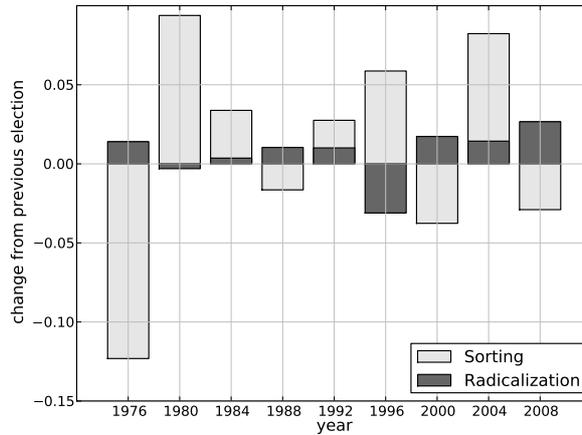


Figure 6: Sorting and radicalization contributions to polarization, 1972-2008

1976. Note that, in those years where both radicalization and sorting increase (1984, 1992, 2004), we draw the effects stacked above each other so that the height of the column in these years is equal to the whole effect (i.e., $\Delta\Psi_t$). In the other years, both radicalization and sorting are drawn starting from zero, and the total change in $\Delta\Psi_t$ is equal to the difference between the positive and the negative column.

It is apparent that sorting is more volatile than radicalization : Sorting increases in five elections, and decreases in four elections, while radicalization increases in most elections, though usually by a small amount. Also, the average absolute change in sorting is considerably larger than the average absolute change in radicalization. This is intuitive because changes in sorting are caused by changes in the distance between the candidates' positions from year to year, and since the candidates change from election to election, while the electorate remains mostly the same as in the previous election, it is very plausible that there are much larger swings possible in sorting than in radicalization.

Our measure of radicalization $\Delta R(t)$ is a specific way of capturing changes in the distribution of voter preferences. Another (essentially model-free) way of measuring radicalization would be to look at the development of the standard deviation of δ and θ in Table 2 in the Appendix.

Obviously, increases in both the standard deviation of δ and of θ translate into positive $\Delta R(t)$. However, there is no clear time trend. The distribution of economic or cultural issue preferences certainly does not appear to become a lot more polarized over time, as this would require a substantial increase in the standard deviations.¹⁷ However, the correlation between economic and cultural conservatism among voters has

¹⁷This result confirms the results of DiMaggio, Evans, and Bryson (1996), Fiorina, Abrams, and Pope (2006) and Fiorina and

increased from a low of 0.18 in 1976 to 0.46 in 2008.¹⁸

These changes in the degree of correlation between δ and θ are what primarily drives the change in our radicalization measure, $\Delta R(t)$. For example, between 1976 and 2004, the standard deviation of δ decreases somewhat, and the standard deviation of θ increases, but also very slightly. However, there is a substantial increase in correlation, so that high δ types are likely to have a low θ , and vice versa; intuitively, this increases the average distance of a voter to the separating line even when the standard deviations remain unchanged. This effect is directly reflected in our measure of radicalization, which shows why $\Delta R(t)$ is a more useful measure than the standard deviation of δ and θ (in addition to having a direct interpretation in the model framework).

5.4 Historical interpretation of our results

We now discuss what the empirical results tell us about the structure of political changes in historical context.

Our time series starts with Nixon vs. McGovern in 1972, a very polarizing election with the Vietnam war and the cultural conflicts it generated as backdrop. Our measures indicate substantial cultural policy divergence and a correspondingly high importance of cultural issues relative to economic ones, k . The next election, pitting Ford versus Carter in 1976, is the low point in our time series regarding policy differences between parties, the relative importance of cultural issues and polarization. It is important to point out that all aspects of voter behavior that our variables measure depend on expectations about the two candidates, rather than on some abstract notion of their “true” positions that we (and the voters) find out in the years after the election. From today’s perspective, Ford was clearly a moderate Republican, and while Jimmy Carter is often considered the most liberal president in the last 50 years, this reputation derives, to a large extent, from his foreign policy positions (say, not bombing Iran in 1979, or being unusually critical of Israel for a U.S. politician), and therefore may have been surprising to voters who probably expected Carter, an evangelical Christian and Southern governor, to be a relatively conservative Democrat.

Ronald Reagan’s success as a conservative in 1980 against Carter is generally considered a key turning point in American politics and initiates a process of ideological realignment of the parties. Cultural policy divergence rebounds to the 1972 level and climbs steadily until plateauing out in 2000. The same features –

Abrams (2008) who all find that overall issue preferences of American voters have remained mostly stable over time.

¹⁸Since high values of δ and low values of θ correspond to cultural and economic conservatism, the negative numbers reported in the last column of Table 2 mean that the two types of conservatism are positively correlated.

a substantial increase in 1980, followed by a steady further increase – is also reflected in the time series of k and of Ψ . Reagan’s “conservative revolution” induces liberal Republicans and conservative Democrats to switch party affiliations throughout the 1980s and 1990s. For example, in 1988, Rick Perry, Norm Coleman, Richard Shelby and David Duke were still Democrats, while Arianna Huffington, Lowell Weicker, Arlen Specter and Lincoln Chafee were still Republicans.¹⁹ When the political elite eventually sort themselves in this way, it reinforces the initial effect of Reagan’s personal conservative policy positions, by making Republicans as a party more socially conservative, and Democrats more socially liberal.

Somewhat surprisingly, the second largest increase in sorting during the whole time period is in 1996 (Dole vs. Clinton) relative to 1992 (Bush I. vs Clinton). As a person, Dole does not necessarily appear to be that much more polarizing than George H.W. Bush. However, it is quite plausible that voter decisions also reflect their perception of the parties at-large, rather than just the specific candidate (essentially, this is the same as when we argued that party switching and the resulting ideological differentiation of parties is reflected in the parties’ respective presidential platforms). In his first term, Bill Clinton attempted a health care reform, and Republicans lead by Newt Gingrich took over the House of Representatives in 1994. Those events may have generated the perception among voters of increased policy divergence between Democrats and Republicans, independent of the specific candidates for the presidency. The wikipedia account of the 1996 campaign (see http://en.wikipedia.org/wiki/United_States_presidential_election,_1996) contains the statement that “Bill Clinton framed the narrative against Dole early, painting him as a mere clone of unpopular House Speaker Newt Gingrich, warning America that Bob Dole would work in concert with the Republican Congress to slash popular social programs, like Medicare and Social Security, dubbed by Clinton as ‘Dole-Gingrich’.” Note that this account is also consistent with the main policy difference in this election being economical rather than cultural, which is reflected in the decrease in k .²⁰

In the 2000 election, Ψ decreases (albeit insignificantly), and increases sharply and significantly in 2004. This is consistent with the narrative among political pundits that George W. Bush had campaigned as a “compassionate conservative” (i.e., a relatively moderate Republican), but that his first term showed that he was much more conservative than expected; moreover, in 2004, he ran against John Kerry, a very liberal Democrat. Thus, policy differences were perceived as relatively small between Bush and Gore in 2000, while the Bush-Kerry election of 2004 was perceived as an election with a stark policy contrast. The

¹⁹See http://en.wikipedia.org/wiki/Party_switching_in_the_United_States.

²⁰1996 is the only election between 1984 and 2000 in which k decreases. The difference between k in 1996 and 2000 is statistically significant on the 95 percent level, while the difference to the 1992 value of k is not.

relative importance of cultural issues, k increases in 2000 and decreases in 2004, primarily driven by a large increase in the perceived economic policy divergence. This is interesting in that most discussions of the 2004 election stress non-economic factors such as the 9/11 terror attacks and the war in Iraq. However, considering that the massive Bush tax cuts were the domestic signature achievement of Bush's first term, the perception that there was a significant increase in economic policy divergence is probably not wrong.²¹

It is also instructive to compare the development of polarization as shown in Figure 5 with different measures of "polarization" in the literature. For example, the percentage of voters casting a straight ticket for President and House (Hetherington 2001, Figure 3), and the percentage of respondents who perceive important differences between the parties (ibid., Figure 5) show a secular increase from the 1970s on, just like Ψ . The same is true of the percentage of strong partisans (Bartels 2000, Figure 1) and the estimated impact of party identification on presidential voting (ibid., Figure 4).²² Overall, this external validation confirms that Ψ measures something that has been interpreted as mass polarization in the existing literature. The main advantage of Ψ is, though, that we can decompose it to disentangle the effects of policy divergence and changes in the voters' preference distribution. Evidently, this would be impossible for any of the measures cited above.

Turning to the development of radicalization, it is interesting to note that, while parties became a lot more differentiated throughout the 1980s and 1990s which lead to a substantial increase in sorting, radicalization changed very little: Figure 6 shows that the radicalization in 1996 is almost exactly the same as in 1976. Thus, the conservative revolution affecting the political elite had arguably very little effect on the preference distribution of the American electorate at-large. Interestingly, this seems to have changed with more substantial increases in radicalization in the last three elections. This may indicate that the elite polarization that started around 1980, apart and in addition to its effect on voter *behavior*, is eventually also having an effect on the fundamental preferred policy positions of the electorate. In Section 5.3, we have argued that the main reason for this is the increase in the correlation between cultural and economic conservatism – voters' ideological preferences become more internally consistent, and it is interesting to note that this coincides temporally with the advent of overtly partisan news channels such as Fox News and MSNBC. Maybe, voters

²¹Also note that our only foreign policy question is quite unspecific (how "warm" a respondent feels toward the U.S. military). In contrast to the Vietnam war in the 1970s, Iraq war opponents always stressed that they opposed the war, not the troops. Thus, the military question we have to use is probably not ideal to pick up the political division on the Iraq war issue.

²²The only substantial qualitative difference is for the 1972 election, which has no particularly remarkable feature in these four measures (and is often measured as less polarizing than 1976), but is identified as a considerably more polarizing election than 1976 by Ψ .

learn from the internally consistent world view that these media outlet provide that cultural conservatives “should” also be economically conservative and vice versa.

By showing the secular increase of k and Ψ , our findings contribute to the controversy about the importance of cultural issues for voters. Journalists and practitioners have long argued that cultural issues are responsible for the secular movement of (economically liberal, but socially conservative) white working class voters from the Democratic party to the Republican party. For example, in the bestseller “What’s the matter with Kansas?”, Thomas Frank (2005) argues that working class voters often vote for Republicans because of cultural issues such as abortion or gay marriage, while their economic interests would be more closely aligned with the Democratic party.²³ This is exactly the party switch that we would expect from culturally conservative voters when k increases and thus the slope of the dividing line increases. On the flip-side, liberal professionals with culturally-liberal, but economically conservative preferences migrate from Republicans to Democrats.

While many pundits consider Frank’s analysis quite compelling, it is based only on qualitative evidence and interpretation. The dominant position of political scientists who have looked for quantitative evidence appears to be rather critical of Frank’s analysis (Bartels 2006, McCarty, Poole, and Rosenthal 2006). However, they do not have a structural model of voter behavior to guide the interpretation of their empirical results. For example, Bartels (2006) analyzes whether working class voters or evangelicals “care more” about social issues such as abortion than other voters, by analyzing whether the probit regression coefficients on social issues are larger. Our model clarifies why this is not a logically sound approach: The size of the probit regression coefficients measuring the importance of different issues for vote choice is determined not by voter preferences, but by party positions, and the parties’ policy difference is the same for all voters. In other words, most evangelicals vote for Republicans not because they “care more about social issues” than the average voter, but rather because their preferred social policies are conservative, and increasingly in line with the Republican platform, and opposed to the Democratic platform. Similarly, McCarty, Poole, and Rosenthal 2006 argue that economic factors have become more, not less important predictors of voting behavior (by comparing their economic variables’ regression coefficients over time). Our results explain why this is so (namely, economic party platforms have diverged), but also indicate that their conclusion that “economic factors have become more, not less important” misses the point that the *relative* importance of cultural factors has increased, because cultural policy platforms diverged even more than economic

²³See also Greenberg (2005) by the pollster who coined the expression “Reagan Democrats” in the 1980s.

platforms.

6 Robustness

6.1 Separating the “cultural” issues

Our measure of “cultural issues” lumps together all “non-economic” policy questions that have been asked continuously in the NES. The advantage of restricting the analysis to two policy dimensions is that it makes it possible to display voter preferences graphically in Figure 3, and it provides for just one marginal rate of substitution between economic issues and all other issues, whose development over time is easier to interpret than the development of $n(n - 1)/2$ different marginal rates of substitution that we get if we instead break policy into n different policy areas.

This said, amalgamating all non-economic policies into one dimension becomes problematic to the extent that policy divergence develops unevenly in different policy areas that are lumped together. For example, suppose that the two parties’ positions on abortion diverged more drastically over time than their military/foreign policy positions. In this situation, aggregating both positions into one “cultural score” implies that we cannot see this change in our results because, by assumption, the relative importance of the different issues for the determination of the cultural preference index δ is fixed.

To analyze the robustness of our results to the aggregation of different cultural issues, we can estimate the model if we treat all cultural questions as separate issues, so that the weights of these issues can change freely between elections. Figure 7 shows that the results for overall polarization are almost identical to the basic model that aggregates all cultural issues. One can also show that the contributions of sorting and radicalization to the change in polarization are also essentially the same as in the basic model, showing the robustness of the results obtained in the basic model with aggregated issues.

6.2 Naive position measurement

As mentioned in the introduction, a direct way of inferring candidates’ positions is to take the answers of NES respondents about the candidates’ positions. As discussed, there are several reasons why this measure could be problematic: First, many respondents might not understand the question about a position in an abstract horizontal policy space (e.g., what positions are really “moderately conservative”?). In con-

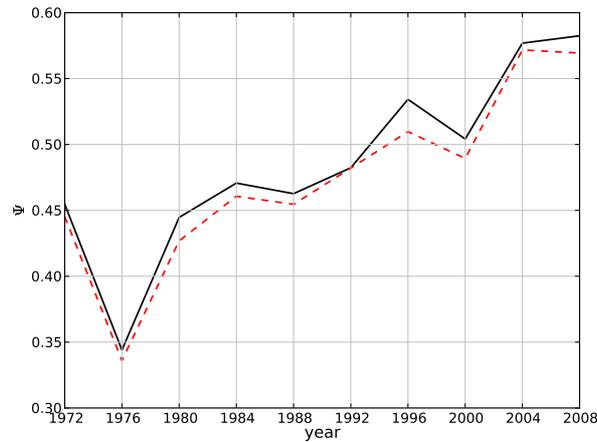


Figure 7: Polarization from 1972 to 2008 when all cultural questions are separate issues (solid line; previous result is dashed line)

trast, the concrete policy questions that we focus on are relatively easy to understand. For the same reason, economists rarely ask consumers directly for their utility function, but rather observe their concrete purchasing decisions, from which they infer the consumers' preferences.

Second, if (what seems reasonable) respondents form their position assignments by comparing different politicians *at the time*, then intertemporal comparability of this measure is low. For example, a competent respondent of the 1980 survey might think that Ronald Reagan was more conservative than Gerald Ford, and therefore assign Reagan to position 6 (conservative). A 2012 respondent might consider Mitt Romney as more moderate than Rick Santorum or Michelle Bachmann and therefore assign Romney to position 5 (moderately conservative). However, this does not imply that Romney is more moderate than Reagan.

Table 3 in the Appendix contains the average score that voters ascribe to the Republican and Democratic candidate in the different elections, as well as the “policy difference” calculated by taking the difference between the scores. Clearly, this produces results that are quite inconsistent with a conventional view of history. For example, Ford was almost exactly as conservative as Nixon, Reagan in 1980 was as conservative as Bush in 2004, and the most conservative Democrat in the last 40 years was Jimmy Carter in 1980.

There is also no statistically significant time trend in the average position of the Democrat or Republican. Thus, this naive measure of political positions does not pick up any significant political polarization trend over the last generation, and does not find any evidence for a “conservative revolution” among Republicans after 1980. Our interpretation is that this absence of evidence is a manifestation of the method's theoretical

problems described above, rather than true evidence of absence. In addition, the fact that Democratic and Republican voters have different views about the candidates' ideological position indicates that absent a model that explains these difference in assessments, these data cannot be used to understand platform differentiation. Finally, it is interesting to note that the correlation between a respondent's placement of the two candidates has become more negative. This may indicate some confirmation about a perceived divergence of platforms, however, it may also indicate that more people became more partisan and place the candidates further apart on the left-right spectrum. Again, absent a model we cannot get clear answers from the data.

6.3 Implications for the theoretical modeling of policy divergence

In this paper, we remain completely agnostic as to what determines the parties' policy choice — the focus of the paper is not to determine what drives policy divergence, but rather to provide the theoretical foundations of a method to measure it. Yet, our analysis can potentially inform the theoretical modeling of candidate competition models.

Our empirical results indicate that the increase in policy divergence precedes any significant radicalization of the electorate by about 20 years. This would indicate that models in which the electoral preference distribution play the determinant role for where candidates will choose their position, such as entry deterrence models, cannot account for the observed change in policy divergence.²⁴

In policy-motivation models, divergence arises from the interplay of policy motivated candidates and uncertainty about the preferences of the median voter. It is hard to argue that our knowledge about the voter preference distribution has decreased in the last generation, especially considering the finding that the voter preference distribution was relatively stable. In contrast, it is quite plausible that there were changes to the extent that office motivation and policy motivation determine the choice of candidates. The modern Presidential primary system was designed to shift power from party elites to ordinary party members (i.e., primary voters) in order for the process to become more democratic. In the Democratic party, this institutional change took effect starting in 1972.²⁵ In both parties the number of states that hold primaries increases throughout the 1970s and 1980s, giving more voice and participation to regular party members at the expense of party elites.²⁶ It is likely that party elites are more interested in winning *per se* (as a winning

²⁴Of course, these models still provide a valuable intuition for some potential reasons of policy divergence.

²⁵See http://en.wikipedia.org/wiki/Democratic_National_Convention.

²⁶For example, just 28 states had primaries or caucuses in the 1976 nomination process, while in 1996, 47 states did.

presidential candidate means that there are a number of executive positions to be distributed), while ordinary party members are primarily policy motivated as none of the office benefits accrues directly to them. When the agents that get to choose the position of the party in the election become more policy-motivated and less office-motivated, this leads to more policy divergence.

Moreover, this process can self-reinforce itself over time: As long as both parties' leaders select very moderate candidates, there is a very limited incentive for regular voters to choose their party membership based on their ideology. But once Republican candidates become more conservative, and Democratic candidates become more liberal, voter registration may follow to become more partisan. Thus, the median regular Republican party member becomes more conservative and the median regular Democrat becomes more liberal. When they select the next candidate, this change in the preferences of primary voters will again be reflected in the candidates that they choose. Since party registration switching is likely to be a rather slow process, it is not implausible that the structural changes in the primary process translate into policy divergence gradually rather than immediately.

7 Conclusion

In this paper, we propose and apply a simple structural model of elections in which voter behavior reflects the extent and direction of party platform divergence. If candidate positions are very similar, then voters choose candidates primarily based on non-policy attributes, whereas if candidate positions are very polarized, then most voters will vote primarily according to their policy preferences. This is true *even if voters care a lot about policy* — without meaningful policy differences between candidates, voters cannot express the direction or intensity of their policy preferences through the act of voting for one of the candidates. In contrast, policy divergence generates a starker choice for voters, and one that is influenced more by the voters' ideal positions relative to the candidates.

Our core contribution is methodological, namely showing that we can recover both the extent and the direction of policy divergence from the voting behavior of voters with different ideal positions. However, the application to NES data is also of substantive interest because it helps to address some of the central questions in American politics. Do voters increasingly split along ideological lines, and if so, what does their behavior tell us about the underlying fundamental causes? Are voters today more polarized than they were a generation ago, or do they just *appear* more polarized because they face more polarizing choices?

These questions can only be answered with the guidance provided by a structural model because one needs to separate the causal effects of changes in voter preferences and changes induced by the political parties.

Our methods are, of course, applicable to other data sets and the questions of policy divergence and polarization in other countries. In particular, it would be very interesting to analyze whether the developments that we identified for the US in the last generation – policy divergence between parties, and stronger divergence on cultural issues than on economic ones – are also reflected in other countries (and in other voting systems such as proportional representation), or whether the experience in the United States is unique in this respect. Such a cross-country comparison will be instrumental for finding out the root cause for the development – *why* is it that parties have diverged over the last generation? And, is this a bad development that should be corrected (and, if so, how?), or is the increased extent of *choice* between parties actually a desirable feature. Evidently, these fundamental questions will require a lot more work, but we hope that the instruments that we have developed in this paper will prove useful in this long-term project.

Appendix

Proof of Theorem 1

Let N_Λ be the set of all i with $\tilde{\lambda}_i < 0$. Then let $X_i = 1 - \tilde{X}_i$ if $i \in N_\Lambda$, and $X_i = \tilde{X}_i$, otherwise.

Similarly, let N_M be the set of all i with $\tilde{\mu}_i < 0$. Then let $Y_i = 1 - \tilde{Y}_i$ if $i \in N_M$, and $Y_i = \tilde{Y}_i$, otherwise.

Note that $\tilde{\lambda}_i \tilde{X}_i = -\tilde{\lambda}_i(1 - \tilde{X}_i) + \tilde{\lambda}_i$. Thus, for $i \in N_\Lambda$ we get $\tilde{\lambda}_i \tilde{X}_i = \lambda_i X_i \sum_{i=1}^n |\tilde{\lambda}_i| + \tilde{\lambda}_i$. For $i \notin N_\Lambda$ it follows that $\tilde{\lambda}_i \tilde{X}_i = \lambda_i X_i \sum_{i=1}^n |\tilde{\lambda}_i|$. Similarly, $\tilde{\mu}_i \tilde{Y}_i = \mu_i Y_i \sum_{i=1}^m |\tilde{\mu}_i| + \tilde{\mu}_i$ for $i \in N_M$ and $\tilde{\mu}_i \tilde{Y}_i = \mu_i Y_i \sum_{i=1}^m |\tilde{\mu}_i|$ for $i \notin N_M$. Thus,

$$\sum_{i=1}^n \tilde{\lambda}_i \tilde{X}_i = \sum_{i=1}^n \lambda_i X_i \sum_{i=1}^n |\tilde{\lambda}_i| + \sum_{i=1}^n \min\{\tilde{\lambda}_i, 0\}, \text{ and } \sum_{i=1}^m \tilde{\mu}_i \tilde{Y}_i = \sum_{i=1}^m \mu_i Y_i \sum_{i=1}^m |\tilde{\mu}_i| + \sum_{i=1}^m \min\{\tilde{\mu}_i, 0\} \quad (23)$$

Since $\delta = \sum_{i=1}^n \lambda_i X_i$, and $\theta = \sum_{i=1}^m \mu_i Y_i$ equation (23) immediately implies (15).

It remains to prove that the modified model corresponds to the original model.

Note that (19) and (17) imply

$$\frac{a_t}{\sigma_t} = \tilde{a}_t - (1 + \rho_t) \sum_{i=1}^m \min\{\tilde{\mu}_i, 0\} + (1 + \alpha_t) \sum_{i=1}^n \min\{\tilde{\lambda}_i, 0\}. \quad (24)$$

(23) implies

$$(1 + \alpha_t) \sum_{i=1}^n \tilde{\lambda}_i \tilde{X}_i = (1 + \alpha_t) \sum_{i=1}^n \lambda_i X_i \sum_{i=1}^n |\tilde{\lambda}_i| + (1 + \alpha_t) \sum_{i \in N_\Lambda} \tilde{\lambda}_i = \frac{k_t}{\sigma_t} \sum_{i=1}^n \lambda_i X_i + (1 + \alpha_t) \sum_{i=1}^n \min\{\tilde{\lambda}_i, 0\} \quad (25)$$

$$(1 + \rho_t) \sum_{i=1}^m \tilde{\mu}_i \tilde{Y}_i = (1 + \rho_t) \sum_{i=1}^m \mu_i Y_i \sum_{i=1}^m |\tilde{\mu}_i| + (1 + \rho_t) \sum_{i \in N_M} \tilde{\mu}_i = \frac{1}{\sigma_t} \sum_{i=1}^m \mu_i Y_i + (1 + \rho_t) \sum_{i=1}^m \min\{\tilde{\mu}_i, 0\}. \quad (26)$$

Next, note that $\sum_{t=1}^s D_t = 1$, since the year dummy for exactly one of the years is 1, and all other ones are zero. Thus, $(1 + \sum_{t=1}^s D_t \alpha_t) = \sum_{t=1}^s D_t (1 + \alpha_t)$. Similarly, it follows that $\sum_{t=1}^s D_t \frac{k_t}{\sigma_t} = \sum_{t=1}^s \frac{D_t}{\sigma_t} \sum_{t=1}^s D_t k_t$. Let $\alpha_0 = \rho_0 = 0$. Then, (25), and (26) imply

$$\begin{aligned} & \left(1 + \sum_{t=2}^s \alpha_t D_t\right) \sum_{i=1}^n \tilde{\lambda}_i \tilde{X}_i - \left(1 + \sum_{t=2}^s \rho_t D_t\right) \sum_{i=1}^m \tilde{\mu}_i \tilde{Y}_i + \sum_{t=1}^s \tilde{a}_t D_t \\ &= \sum_{t=1}^s D_t (1 + \alpha_t) \sum_{i=1}^n \tilde{\lambda}_i \tilde{X}_i - \sum_{t=1}^s D_t (1 + \rho_t) \sum_{i=1}^m \tilde{\mu}_i \tilde{Y}_i + \sum_{t=1}^s \tilde{a}_t D_t \end{aligned}$$

$$\begin{aligned}
&= \sum_{t=1}^s D_t \frac{k_t}{\sigma_t} \sum_{i=1}^n \lambda_i X_i - \sum_{t=1}^s \frac{D_t}{\sigma_t} \sum_{i=1}^m \mu_i Y_i \\
&\quad + \sum_{t=1}^s D_t \left[\tilde{a}_t + (1 + \alpha_t) \sum_{i=1}^n \min\{\tilde{\lambda}_i, 0\} - (1 + \rho_t) \sum_{i=1}^m \min\{\tilde{\mu}_i, 0\} \right] \\
&= \sum_{t=1}^s \frac{D_t}{\sigma_t} \left[\sum_{t=1}^s k_t D_t \sum_{i=1}^n \lambda_i X_i - \sum_{i=1}^m \mu_i Y_i + \sum_{t=1}^s D_t a_t \right],
\end{aligned}$$

where the last equality follows from (24). The two models are therefore equivalent. \blacksquare

The General Model

We generalize the model to the case with J cultural or ideological position $\delta_P = (\delta_{j,P})_{j=1,\dots,J} \in [0, 1]^J$, $P \in \{D, R\}$, for candidates $P = D, R$ in addition to the economic position g_P . Thus, voter's utility from candidate P is given by

$$u(\delta, \theta, \xi_P) = \theta v(g_P) - c_P - \sum_{j=1}^J w_j (\delta_j - \delta_{j,P})^2 + \xi_P, \quad (27)$$

where $w_j > 0$, $j = 1, \dots, J$ are weights on the relative importance of the individual issues. For standard Euclidean preferences, all weights would be identical.

It is easy to verify that (3) generalizes to

$$\theta(\delta, \xi, g_D, g_R) = \frac{2 \sum_{j=1}^J \delta_j w_j (\delta_{j,R} - \delta_{j,D}) + c_D - c_R - \sum_{j=1}^J w_j (\delta_{j,R}^2 - \delta_{j,D}^2) + \xi}{v(g_D) - v(g_R)}. \quad (28)$$

Let

$$k_j = \frac{2w_j(\delta_{j,R} - \delta_{j,D})}{v(g_D) - v(g_R)}, \quad a = \frac{c_D(g_D) - c_R(g_R) - \sum_{j=1}^J w_j(\delta_R^2 - \delta_D^2) + \bar{\xi}}{v(g_D) - v(g_R)}, \quad (29)$$

where $\bar{\xi} = E[\xi]$. Let $k = (k_j)_{j=1,\dots,J}$. Then (28) implies that the separating hyperplane is given by $\theta = k \cdot \delta + a$, and a citizen votes Republican if and only if

$$\theta - k \cdot \delta - a - \varepsilon < 0. \quad (30)$$

Let $X_{i,j}$, $i = 1, \dots, n_j$, $j = 1, \dots, J$ and Y_i , $i = 1, \dots, m$ be random variables that describe the answers to survey questions on issues $j = 1, \dots, J$ and economic issues, respectively. Let $\delta_j = \sum_{i=1}^{n_j} \lambda_{i,j} X_{i,j}$ and $\theta = \sum_{i=1}^m \mu_i Y_i$, where, $\lambda_{i,j}$ and μ_i are parameters to be estimated.

We use the same normalizations for $X_{i,j}$ and $Y_{i,j}$ as in the two-dimensional case, i.e., the lowest and highest realizations for each question are 0 and 1; and high values on $X_{i,j}$ and $Y_{i,j}$ increase the estimated value of δ_j and θ , respectively. Similarly, we again normalize $\sum_{i=1}^{n_j} \lambda_{i,j} = 1$ for all $j = 1, \dots, J$ and $\sum_{i=1}^m \mu_i = 1$.

Let D_t , $t = 1, \dots, s$ be the year dummy for year $t = 1, \dots, s$ (i.e., $D_t = 1$ if the observation occurred in year t , and 0 otherwise). Then (11) generalizes to

$$\Phi \left(\sum_{t=1}^s \frac{D_t}{\sigma_t} \left[\sum_{j=1}^J \left(\sum_{i=1}^s D_t k_{j,t} \right) \left(\sum_{i=1}^n \lambda_{i,j} X_{i,j} \right) - \sum_{i=1}^m \mu_i Y_i + \sum_{t=1}^s D_t a_t \right] \right). \quad (31)$$

Again, we consider the model without restriction on the $\tilde{\lambda}_i$, and $\tilde{\mu}_i$ and where $\tilde{X}_{i,j}$ and \tilde{Y}_i are the observations only normalized to be between 0 and 1. The probability of voting Republican is given by

$$\Phi \left(\sum_{j=1}^J \left(1 + \sum_{t=2}^s \alpha_{j,t} D_t \right) \left(\sum_{i=1}^n \tilde{\lambda}_{i,j} \tilde{X}_{i,j} \right) - \left(1 + \sum_{t=2}^s \rho_t D_t \right) \left(\sum_{i=1}^m \tilde{\mu}_i \tilde{Y}_i \right) + \sum_{t=1}^s \tilde{a}_t D_t \right), \quad (32)$$

Denote by $d_{t,\ell}$, $\tilde{x}_{i,j,\ell}$, and $\tilde{y}_{i,\ell}$ observation ℓ of random variables D_t , $\tilde{X}_{i,j}$, and \tilde{Y}_i , respectively. Let

$$z_\ell = \sum_{\ell=1}^L \sum_{j=1}^J \left(1 + \sum_{t=2}^s \alpha_{j,t} d_{t,\ell} \right) \left(\sum_{i=1}^n \tilde{\lambda}_{i,j} \tilde{x}_{i,j,\ell} \right) - \left(1 + \sum_{t=2}^s \rho_t d_{t,\ell} \right) \left(\sum_{i=1}^m \tilde{\mu}_i \tilde{y}_{i,\ell} \right) + \sum_{t=1}^s \tilde{a}_t d_{t,j}, \quad (33)$$

and let $v_j = 1$ if the voter in observation j votes Republican, and $v_j = 0$ if he votes Democrat. To estimate α_i , β_i , $\tilde{\lambda}_i$, $\tilde{\mu}_i$, and \tilde{a}_i , we maximize the log-likelihood function, i.e., solve

$$\max_{\{\alpha_i, \rho_i | i=2, \dots, s\}, \{\tilde{a}_i | i=1, \dots, s\}, \{\tilde{\lambda}_{i,j} | i=1, \dots, n, j=1, \dots, J\}, \{\tilde{\mu}_i | i=1, \dots, m\}} \sum_{j=1}^J v_j \ln \Phi(z_j) + (1 - v_j) \ln (1 - \Phi(z_j)). \quad (34)$$

Theorem 1 immediately generalizes in the obvious way. For example, the definition of θ in (11) remains unchanged, and in the definition of δ we only need to replace δ by δ_j and X_i by $X_{i,j}$. Similarly, in (18) we replace k_t by $k_{j,t}$ and $\tilde{\lambda}_i$ by $\tilde{\lambda}_{i,j}$.

Tables

| | 1972–1988 | 1992–2008 | $(\lambda, \mu)_{1972-1988}$ | $(\lambda, \mu)_{1992-2008}$ |
|--|---------------------------|---------------------------|------------------------------|------------------------------|
| military (thermometer) | 1.158 [0.798,1.577] | 1.200 [0.874,1.628] | 0.329 [0.264,0.388] | 0.305 [0.246,0.364] |
| aid to minorities (high answers = against aid to minorities) | 0.551 [0.345,0.766] | 0.633 [0.437,0.859] | 0.157 [0.097,0.216] | 0.161 [0.110,0.212] |
| black (thermometer) | -1.374 [-1.845,-0.987] | -0.982 [-1.377,-0.671] | 0.390 [0.321,0.454] | 0.250 [0.190,0.307] |
| role of women (high answers = women’s place is in the house) | 0.055 [-0.110,0.226] | 0.318 [0.129,0.529] | 0.016 [0.001,0.063] | 0.081 [0.034,0.127] |
| abortion (high answers = should be legal) | -0.267 [-0.433,-0.113] | -0.698 [-0.915,-0.522] | 0.076 [0.032,0.118] | 0.177 [0.138,0.220] |
| attends church | 0.117 [0.014,0.230] | 0.105 [0.008,0.203] | 0.033 [0.005,0.061] | 0.027 [0.003,0.051] |
| big business (thermometer) | -1.537 [-1.907,-1.235] | -0.860 [-1.183,-0.584] | 0.366 [0.325,0.407] | 0.235 [0.176,0.288] |
| union (thermometer) | 1.588 [1.296,1.934] | 1.806 [1.480,2.189] | 0.378 [0.340,0.417] | 0.494 [0.444,0.546] |
| government standard of living (high answer = no gov. welfare) | -1.078 [-1.321,-0.865] | -0.987 [-1.239,-0.770] | 0.256 [0.219,0.295] | 0.270 [0.224,0.319] |

Table 1: Estimation of Parameters; 95 percent confidence interval

| year | average δ | std. dev. δ | average θ | std. dev. θ | correlation |
|------|------------------|--------------------|------------------|--------------------|-------------|
| 1972 | 0.499 | 0.147 | 0.502 | 0.159 | -0.237 |
| 1976 | 0.504 | 0.139 | 0.453 | 0.168 | -0.183 |
| 1980 | 0.502 | 0.132 | 0.489 | 0.165 | -0.284 |
| 1984 | 0.472 | 0.138 | 0.501 | 0.169 | -0.260 |
| 1988 | 0.497 | 0.131 | 0.480 | 0.173 | -0.269 |
| 1992 | 0.474 | 0.141 | 0.487 | 0.165 | -0.322 |
| 1996 | 0.494 | 0.127 | 0.473 | 0.160 | -0.327 |
| 2000 | 0.497 | 0.127 | 0.477 | 0.164 | -0.340 |
| 2004 | 0.497 | 0.138 | 0.510 | 0.171 | -0.396 |
| 2008 | 0.486 | 0.140 | 0.535 | 0.183 | -0.458 |

Table 2: Cultural and economic indices: Average and standard deviation

| year | av. Dem. | av. Rep. | av. Dem -av. Rep. | Dem. view Dem. | Dem. view Rep. | Rep. view Dem. | Rep. view Rep. | corr. |
|------|----------|----------|----------------------|-------------------|-------------------|-------------------|-------------------|--------|
| 1972 | 2.443 | 4.928 | 2.485 | 2.634 | 5.225 | 2.188 | 4.850 | -0.174 |
| 1976 | 3.222 | 4.926 | 1.704 | 3.359 | 5.161 | 2.848 | 4.883 | -0.186 |
| 1980 | 3.711 | 5.220 | 1.509 | 4.014 | 5.402 | 3.288 | 5.239 | -0.329 |
| 1984 | 3.447 | 4.986 | 1.539 | 3.665 | 5.186 | 3.050 | 5.184 | -0.333 |
| 1988 | 3.211 | 5.137 | 1.926 | 3.649 | 5.253 | 2.449 | 5.331 | -0.318 |
| 1992 | 3.173 | 5.080 | 1.906 | 3.437 | 5.280 | 2.483 | 5.227 | -0.247 |
| 1996 | 3.114 | 5.155 | 2.042 | 3.599 | 5.157 | 2.146 | 5.396 | -0.354 |
| 2000 | 3.159 | 5.029 | 1.870 | 3.431 | 5.162 | 2.571 | 5.150 | -0.269 |
| 2004 | 2.955 | 5.224 | 2.269 | 3.393 | 5.358 | 2.339 | 5.402 | -0.393 |
| 2008 | 3.272 | 4.861 | 1.588 | 3.579 | 4.870 | 2.067 | 5.238 | -0.444 |

Table 3: Naive Placement of Candidate Postions

References

- Ansolabehere, S., J. Rodden, and J. Snyder (2006). Purple America. *Journal of Economic Perspectives* 20(2), 97–118.
- Banks, J. S. and J. Duggan (2005). Probabilistic voting in the spatial model of elections: The theory of office-motivated candidates. In *Social Choice and Strategic Decisions*, pp. 15–56. Springer: Berlin, Heidelberg.
- Bartels, L. (2000). Partisanship and voting behavior, 1952-1996. *American Journal of Political Science* 44(1), 35–50.
- Bartels, L. (2006). What’s the matter with ‘What’s the matter with Kansas?’ . *Quarterly Journal of Political Science* 1(2), 201–26.
- Bernhardt, D., J. Duggan, and F. Squintani (2006). Private polling in elections and voter welfare. mimeo, University of Illinois.
- Callander, S. (2005). Electoral competition in heterogeneous districts. *Journal of Political Economy* 113, 1116–1145.
- Callander, S. (2008). Political motivations. *Review of Economic Studies* 75, 671–697.
- Calvert, R. L. (1985). Robustness of the multidimensional voting model: Candidate motivations, uncertainty, and convergence. *American Journal of Political Science* 29, 69–95.
- Castanheira, M. (2003). Why vote for losers? *Journal of the European Economic Association* 1(5), 1207–1238.
- DiMaggio, P., J. Evans, and B. Bryson (1996). Have American’s social attitudes become more polarized? *American Journal of Sociology* 102, 690–755.
- Dixit, A. and J. Londregan (1995). Redistributive politics and economic efficiency. *American Political Science Review* 89(4), 856–866.
- Downs, A. (1957). *An economic theory of democracy*. Harper.

- Fiorina, M. and S. Abrams (2008). Political polarization in the American public. *Annual Review of Political Science* 11(1), 563–588.
- Fiorina, M., S. Abrams, and J. Pope (2006). *Culture war? The myth of a polarized America*. Pearson Longman.
- Frank, T. (2005). *What's the matter with Kansas?: How conservatives won the heart of America*. Holt Paperbacks.
- Greenberg, S. (2005). *The two Americas: Our current political deadlock and how to break it*. St. Martin's Griffin.
- Gul, F. and W. Pesendorfer (2009). Partisan politics and aggregation failure with ignorant voters. *Journal of Economic Theory* 144, 146–174.
- Hetherington, M. J. (2001). Resurgent mass partisanship: The role of elite polarization. *American Political Science Review* 95(3), 619–631.
- Hinich, M. (1978). The mean versus the median in spatial voting games. In P. Ordeshook (Ed.), *Game Theory and Political Science New York*. NYU Press: New York.
- Krasa, S. and M. Polborn (2010). Competition between specialized candidates. *American Political Science Review* 104(4), 745–765.
- Krasa, S. and M. Polborn (2011). Social ideology and taxes in a differentiated candidates framework. University of Illinois.
- Lindbeck, A. and J. Weibull (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice* 52, 273–297.
- Martinelli, C. (2001). Elections with privately informed parties and voters. *Public Choice* 108(1), 147–167.
- McCarty, N., K. Poole, and H. Rosenthal (2006). *Polarized America: The dance of ideology and unequal riches*. MIT Press.
- Palfrey, T. R. (1984). Spatial equilibrium with entry. *Review of Economic Studies* 51, 139–156.
- Poole, K. and H. Rosenthal (1984). The polarization of American politics. *Journal of Politics* 46(04), 1061–1079.
- Poole, K. and H. Rosenthal (1985). A spatial model for legislative roll call analysis. *American Journal of Political Science* 29(2), 357–384.
- Poole, K. and H. Rosenthal (2000). *Congress: A political-economic history of roll call voting*. Oxford University Press, USA.
- Soubeyran, R. (2009). Does a disadvantaged candidate choose an extremist position? *Annales d'Economie et de Statistique*, forthcoming.
- Wittman, D. (1983). Candidate motivation: A synthesis of alternative theories. *American Political Science Review* 77, 142–157.