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

Forthcoming, Journal of Law and Economics (2014)

Stefan Krasa† Mattias Polborn‡

February 11, 2014

Abstract

One of the most widely discussed phenomena in American politics today is the perceived increasing partisan divide that splits the U.S. electorate. A contested question is whether this diagnosis is actually true, and if so, what is the underlying cause.

We propose a new method that simultaneously estimates voter preferences and parties' positions on economic and "cultural" issues. We apply the model to U.S. presidential elections between 1972 and 2008. The model recovers candidates' positions from voter behavior, and decomposes changes in the overall political polarization of the electorate into changes in the distribution of voter ideal positions ("voter radicalization") and consequences of elite polarization ("sorting").

*We would like to thank seminar participants at the University of Rochester (Wallis Conference 2011), Northwestern University/University of Chicago (PECA 2012), Princeton University, University of Cologne and University of Bielefeld for helpful comments. Both authors gratefully acknowledge financial support from National Science Foundation Grant SES-1261016. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation or any other organization.

†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.

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

1 Introduction

One of the most widely discussed phenomena in American politics today is the perceived increase in "polarization," among both party elites and voters. Polarization in Congress has increased substantially over the last 30 years, from a historic low achieved between roughly 1940 and 1980 (e.g., Poole and Rosenthal 2000; Groseclose, Levitt and Snyder 1999). Elite polarization also appears to be prevalent among party members and activists (Abramowitz and Saunders 1998, 2008; Harbridge and Malhotra 2011).

In contrast, beliefs about mass polarization vary substantially in the literature. On the one hand, many political 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".[1] On the other hand, research that analyzes voter preferences on different policy issues directly rather than voter behavior finds little evidence that the preferences of the American electorate have moved from moderate positions to more extreme ones over the last generation (e.g. DiMaggio, Evans and Bryson 1996; Fiorina, Abrams and Pope 2006; Bartels 2006; Fiorina and Abrams 2008; Levendusky 2009).

The tension between increasingly partisan voter behavior on the one hand and no fundamental change in voter preferences on the other is puzzling: If voters' fundamental preferences on issues did not change, why do they now *act* in more partisan ways? To answer this question, we need a framework that provides for an explicit mechanism linking party elite actions and mass voting behavior. In this paper, we develop such a model that allows us to answer the following important questions: First, have the masses in fact become more polarized, or is what has been perceived and identified as polarization really just a reflection of changes in elite behavior? Second, to what extent have elites and masses contributed, if at all, to changes in polarization? Third, is polarization driven primarily by economic or by cultural issues?

To gain an intuitive understanding of the effects captured by the model, consider a society in

[1] "On His High Horse," (November 9, 2002) and "America's Angry Election," (January 3, 2004).

which the parties' policy platforms are virtually indistinguishable. In this case, whether Democrats or Republicans win hardly makes a difference for the implemented policy, so that voters may not base their vote choices on their ideological preferences, but rather on their personal and idiosyncratic perceptions of the candidates. Superficially, when outside observers analyze the ideological determinants of voting behavior in this society, it looks as if voters do not care about issues. However, if party elites become more polarized over time, creating a more meaningful choice, then voters will expose previously buried ideological divisions among them, even if their fundamental preferences on policies remain constant: In short, elite polarization can beget voter behavior that *appears* more polarized, but actually is not a reflection of the voter preference distribution becoming more extreme. Moreover, whether voters appear to be more strongly polarized on economic issues or on cultural ones depends crucially on whether the distance between the parties is larger on economic or cultural issues.[2]

Because policy divergence between parties influences how voters' ideal positions on policy issues translate into vote choices, observing voters' behavior allows us to draw inferences about policy divergence. Using NES data from the U.S. presidential elections between 1972 and 2008, we show how we can use observations of voter preferences on different policy issues and voters' choices of which candidate to vote for, to simultaneously estimate the ideal positions of voters on economic and cultural issues, and the difference between Democratic and Republican presidential candidates' positions on those issues during this time period.

In contrast to models that only focus on measuring politicians' positions, our model combines an analysis of politicians and voters, thus providing us with a better understanding of the underlying causes of electoral polarization: Does the electorate look more politically polarized today than 30 years ago, and if so, is party platform divergence, a change in the voters' preferences, or both responsible for this? To analyze these questions, we define a measure of the electorate's

[2]That voters' issue preferences more strongly affect their vote choices, the more distant party positions are from each other, assumes only rational behavior by voters and not changes in their underlying policy preferences. We do not assume that elite polarization on an issue "makes people think more about that issue" and that they consequently *develop* more radical preferences on the issues. Rather, rational voters are always aware of their issue preferences, but they will only condition their vote choice on their issue preferences, if both candidates take different positions on these issues.

polarization on political issues. It quantifies the degree to which voters' candidate choices depend on their preferred issue positions. Our estimation procedure provides a distribution of voters' ideal points and the positions of candidates, in different elections. We can therefore logically 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 "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."

In contrast to existing methods of position estimation that derive politicians' positions through observation of their votes on certain proposed legislation in a legislature, our method measures the policy distance that voters *perceive* between candidates. Our method thus complements these methods because voter-perceived positions are clearly important as well: After all, voters should care about the positions that each candidate *will* take if elected, rather than about his past positions as reflected in his voting record.

There are at least three reasons why focusing purely on past actions may not be a perfect predictor of either the voters' perception or the candidates' future behavior if elected as President: First, the constitutional competences of the President are very different from those of Congress, so a candidate's Congressional voting record may not necessarily be all that relevant for voters. For example, Ron Paul's DW-Nominate score was more conservative than 99 percent of Republican Congressmen. However, in his Presidential nomination runs in 2008 and 2012, he enjoyed considerable support from more moderate voters because of his foreign policy positions that were never reflected in his voting record, because the House of Representatives rarely gets to vote on foreign policy decisions.

Second, the President does not set policy in isolation, but rather in collaboration with other actors from his administration and party. For example, vice presidential candidates are often said to be chosen to provide ideological "balance" to the ticket. But if this is true, then even if a voting-history based concept were to perfectly measure a candidate's own position, we do not know a

priori whether voters in the Presidential election evaluate only the Presidential candidates' own positions, or some amalgamation of their positions and those of their running mates or other actors in their respective parties; for example, in 1996, Bill Clinton rather successfully framed his opponent as "Dole-Gingrich."Finally, politicians' positions may change over time. They may attempt to explicitly disavow positions that they have previously taken (e.g., Mitt Romney and "Romney-care"), and whether voters believe in their new positions or in the position that materializes in their historical vote choices is an empirical issue that is not a priori clear.

In the next section, we discuss some of the related literature. Section 3 sets out our model. In Section 4, we define our key concepts, show how they correspond to the model and provide the theoretical basis for the estimation. In Sections 5, 6 and 8, we apply our methods to National Election Survey data from U.S. Presidential elections between 1972 and 2008. In Section 7, we analyze how increased voter participation would have affected these Presidential elections. Section 9 discusses different issues and concludes. The Online Appendix contains proofs, a generalized model and some robustness analysis.

2 Related literature

Starting with the seminal contribution of Downs (1957), there is a large theoretical literature on platform convergence or divergence in variations of the spatial model.[3] Our empirical results show a substantially stronger policy divergence between parties at the end of our observation period than in the beginning. While we do not propose a theoretical explanation for why this is the case, measurement of policy divergence clearly is an extremely important tool for the evaluation of these theories of party platform choice in electoral competition.

One of the main topics that our model addresses is the notion of political polarization. The usage of the term "polarization" is non-uniform in the literature. Many authors use polarization as

[3]This 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).

synonymous to "policy divergence between parties" (see, e.g., McCarty et al. (2006); Feddersen and Gul (2013)). In contrast, Esteban and Ray (1994, 1999, 2011); Duclos et al. (2004) define polarization as a property of the preference distribution of voters; specifically, in their definition, polarization captures the notion of a society consisting of different groups in which voters in each group are very similar to each other, but very dissimilar to voters in other groups. Our notion of *polarization* captures the interaction of two underlying forces: Preference diversity among voters, and party policy divergence that creates an outlet for the expression of this preference diversity. Furthermore, we can measure the respective contributions of these two forces to polarization as radicalization and sorting, respectively.

We provide a new method of comparing policy divergence over time. The standard method of determining the positions of politicians is based on the seminal work of Poole and Rosenthal (1984, 1985) that we discuss in more detail in Section 9.1. This DW-Nominate method relies on the analysis of many votes by the politicians in legislatures and therefore runs into problems when evaluating candidates who have not served in the same legislature (see Section 9.1). Furthermore, by explicitly distinguishing between economic and cultural issues, our method can provide information on the temporal development of policy divergence in different areas of policy, something that the DW-Nominate method is not designed to deliver.

In a one-dimensional framework, Degan (2007) estimates a distribution of voter ideal positions and candidate valences for the 1968 and 1972 U.S. Presidential elections, assuming that candidate positions are given by their respective DW-Nominate scores in the Senate. In contrast, our method allows for a simultaneous estimation of voter ideal points and candidate positions and can be applied over much longer time periods.

One core intuition behind our structural model is present as a qualitative idea in Fiorina et al. (2006) who point out that, in a multidimensional setting, the direction of elite polarization influences the direction of the fault line through the electorate, and that this effect constitutes a severe challenge for empirical studies that analyze the determinants of voter behavior. They correctly recognize that interpreting the size of regression coefficients as equivalent to the "importance" of the corresponding question for voters is not logically correct, and conclude (p. 183): "The findings

of scores if not hundreds of electoral studies are ambiguous. The problem most deeply afflicts attempts to study electoral change by conducting successive cross-sectional analyses and comparing the results." However, they do not use this insight positively to develop it into a structural model, and this is our fundamental contribution.

Our analysis also contributes to an important substantive debate in the literature about what type of issues – economic or cultural – drive vote choice today, and how their relative effects might have changed over time. A common impression among political journalists and practitioners is that moral issues have become more important for defining the parties and their supporters. For example, in the popular bestseller "What's the matter with Kansas?", Frank (2005) argues that poor people 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. Hunter (1992), Shogan (2002) and Greenberg (2005) present similar "culture-war" arguments. However, many political scientists challenge this thesis, and emphasize the importance of economic issues in explaining voter preferences for candidates (e.g., Bartels (2006); McCarty et al. (2006); Gelman et al. (2008); Bartels (2010)). Ansolabehere, Rodden and Snyder (2006) provide some mixed evidence, and show a substantially increased importance of moral issues for vote choices in the 1990s relative to the 1970s and 80s, but also find that economic factors are still more important for voters than purely moral ones. Our main contribution to this literature is that we provide a structural model in which we can analyze the *relative* importance of economic and cultural factors for vote choices, as well as the underlying reasons for the shift towards a higher importance of cultural issues.

3 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 $\theta \in [0,1]$ measuring

his preferences for public goods, and a parameter $\xi_P \in \mathbb{R}$ measuring the impact of the personal charisma of the candidate $P = D, R$ on the voter. A voter's utility from candidate P is given by

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

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, the model has one economic and one cultural dimension, but in the Appendix, we describe how it can be modified for an arbitrary number of ideological issues. Also, our focus is on analyzing the consequences of policy divergence for voter behavior. 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 (2014) 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.

[4] We 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.

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

[6] Note 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.

4 Analysis of the Model

4.1 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

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

We assume that the Democrat provides weakly 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

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

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 conservative, 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.

[7] From 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 position. These normalizations make sense in the U.S. context.

4.2 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

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

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

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

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

$$\theta - k\delta - a - \varepsilon < 0. \tag{6}$$

Let X_i, $i = 1, \ldots, n$ and Y_i, $i = 1, \ldots, m$ 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

[8] Clearly, 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.

[9] In 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.

(6) implies that the probability that a voter votes Republican is given by

$$\Phi\left(\frac{1}{\sigma}\left[k\sum_{i=1}^{n}\lambda_i X_i - \sum_{i=1}^{m}\mu_i Y_i + a\right]\right). \tag{7}$$

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

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

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] In Section 8.2, we discuss how to account for changes of σ_ξ over time. In Section 8.3, we show how to account for misspecification of σ_ξ because of missing questions in the surveys.

Using (4) implies

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

We can use equations (17) and (18) in Theorem 1 (in Section 4.4 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.

[10] In 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. Likewise, the DW-Nominate method assumes that the error term is "constant across all of American history" (Poole and Rosenthal, 2011, p. 27). See our discussion in Section 8.

4.3 Polarization, Radicalization and Sorting

"Polarization" is a central issue in the analysis of American political behavior. As mentioned in the introduction, many commentators diagnose a sharp and increasing partisan divide that splits the U.S. electorate, but there is no general agreement on a formal definition of what constitutes "polarization." Intuitively, it does not make sense to define polarization by how close 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. 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.

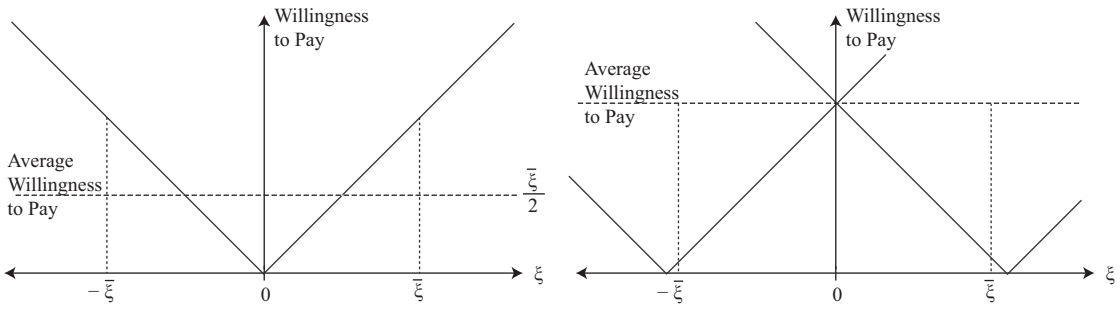

Figure 1: Willingness to pay with policy convergence and policy divergence

Consider Figure 1. Suppose there are two "groups" (distinguished by, e.g., ideological positions, gender, race, ethnicity) with different policy preferences, and each individual also receives

an idiosyncratic preference shock, on top of their policy preference, both measured in terms of willingness to pay. In the left panel, both ideological groups receive the same policy payoff from both candidates, so the average individual willingness to pay for a victory of the preferred candidate is only based on the individual preference shock and is equal to $\int |\xi| d\Phi_\xi$, where Φ_ξ is the distribution of ξ, centered around 0. In Figure 1, this distribution is uniform between $-\bar{\xi}$ and $\bar{\xi}$, and the average willingness to pay is $\bar{\xi}/2$.

In the right panel, the two groups receive different policy payoffs from the two candidates. As drawn, idiosyncratic preferences never completely offset the voters' policy preferences. Thus, the average willingness to pay for a victory of one's preferred candidate is equal to the absolute value of the policy preference, and so is larger in the right panel, and it is the larger, the larger are the policy preferences of the two groups.

Of course, the willingness to pay concept of polarization cannot be operationalized directly because there are no opinion polls that ask voters for their willingness to pay. However, both increased policy divergence between candidates and increasingly polarized ideological preference do have similar observable implications for *voter behavior* as they would for a willingness-to-pay measure. In the left panel where the two groups have no systematic policy preferences among the candidates, observed voting behavior does not differ between the two groups – knowing a voter's group membership is not informative about the individual's vote choice. In contrast, in the right panel, observing an individual's group membership *is* informative about the individual's vote choice, and is the more informative, the stronger the difference between the groups' policy payoffs.

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 voting behavior. In our application, we are interested in the *ideological* polarization of the U.S. electorate, and its change over time.[11] That is, we will construct a measure of how much do voters divide along their observable ideological positions. In the following, we will skip "ideological" when no confusion

[11]For other applications, one can in principle focus on other types of demographic polarization, such as gender, racial, ethnic or religious polarization that tell us how the electorate splits along the lines defined by these characteristics. Also, our measure of polarization will allow us to make statements like "society is more racially polarized than economically polarized", or vice versa; see the discussion of Figure 6 in Section 6.3 below.

can arise and just talk about polarization.

To formalize our concept of polarization, suppose that we have to predict the voting behavior of a large group of voters in a close 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 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, if 70 percent of voters vote for the Republican candidate in an election, 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:

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

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 $2/I$ 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 voter is either conservative or liberal, and every conservative votes Republican, and every liberal votes Democratic. (Most) voters know which party they will vote for before they know the valence of the actual candidates – 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 change).

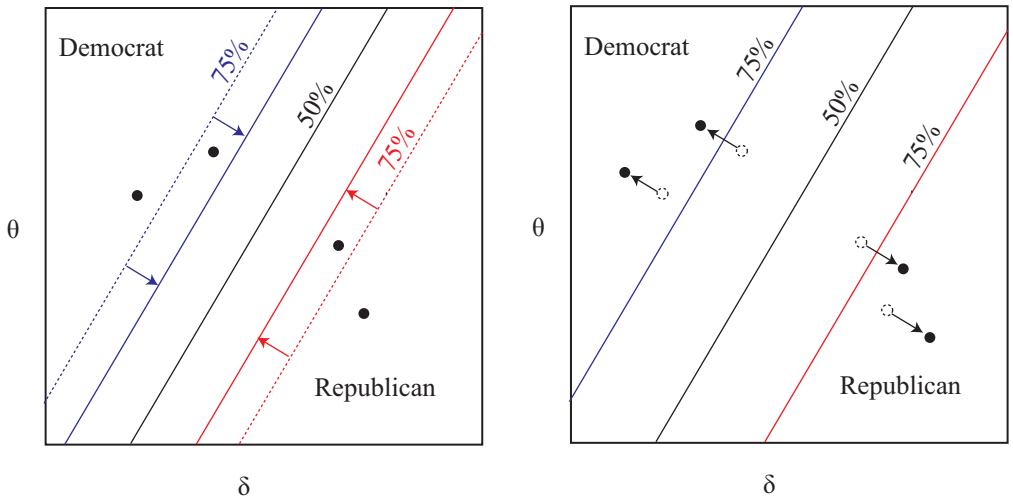

Figure 2: Increasing Polarization through Sorting and Radicalization

Figure 2 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. The distance from the 50% line to any other "isoprobability" line, such as the 75% line in the graph, is proportional to $\sigma/\sqrt{1+k^2}$. Thus, (4) and (8) imply that the distance is proportional to $\dfrac{\sigma_\xi}{\sqrt{[v(g_D)-v(g_R)]^2+4(\delta_R-\delta_D)^2}}$. As a consequence, increased policy divergence moves the isoprobability lines closer together in the left panel of Figure 2, 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 2 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 $\Delta\Psi_t = \Psi(t,t) - \Psi(t-4, t-4)$. When we keep the electorate of the last election in $t-4$ fixed and only vary the politicians' positions, we obtain $\Delta S(t) \equiv \Psi(t-4, t) - \Psi(t-4, t-4)$, the level of sorting in year t. 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 a hypothetical 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.

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 6.3), 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."

4.4 Estimation Procedure

To determine 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 for k and σ to change over time and thus index them by the year of the election. Let D_t be the year dummy (i.e., $D_t = 1$ if the observation occurred in year t, and 0 otherwise). Then (7) generalizes to

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

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

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

where there are no restrictions on $\tilde{\lambda}_i$ 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 not requiring that higher realizations of the response to each question increase δ and θ.

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

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

and let $v_\ell = 1$ if the voter in observation ℓ votes Republican, and $v_\ell = 0$ 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

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

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] 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* $t \in \{1, \ldots, s\}$; $\tilde{\lambda}_i$, $i \in \{1, \ldots, n\}$; $\tilde{\mu}_i$, $i \in \{1, \ldots, m\}$, *be the parameters of the modified model in* (12). *Then the parameters of the original model* (11) *are determined as follows:*

[12] We 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. The computer code for performing the estimation can be obtained from the authors.

1. δ and θ are given by

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

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

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

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

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

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

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

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

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

5 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 standard opinion polls or exit 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]

[13] Because we need continuously available questions, we start our analysis in 1972: Moving to the 1960s would have meant losing a substantial number of questions, while moving the start date into the late 1980s would have expanded the number of questions for which data are available, but at the cost of shortening the time series substantially.

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.[15]

[14] Also, 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.

[15] Data on respondent's demographic characteristics (such as gender and race) is available, but we prefer not to use these variables as "policy positions," as the NES has information on policy preferences. In section 8.3 we show that our results also hold if the questions in the survey are only imperfectly correlated with the actual policy issues in the different elections, and if some relevant questions are missing.

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

We ignore the respondents' partisan affiliation and self-placement on a liberal-to-conservative scale, because including such a measure would defy the purpose of our analysis. First, 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 NES respondents placed Obama strictly to the right of McCain in 2008. Second, 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.

6 Empirical Results

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

We first find the weights of different issue questions for the determination of the voters' ideological positions. As described in Section 4.4, 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.[16] 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 voter behavior many years ago, and what made a person economically or culturally

both economic and non-economic policy preferences, this would not tell us anything definitive about the policy area in which the parties diverged. Also, 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 insignificant.

[16] If we were to choose just one year as the base period, then the modified model of (12) specifies a standard probit model. However, we still need Theorem 1 to retrieve the actual model parameters.

conservative in the 1970s may be different from 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.

Issue	military (thermometer)	aid to minorities	black (thermometer)	role of women	abortion	attends church
$\lambda_{1992-2008}$	0.305	0.161	0.250	0.081	0.177	0.027
conf. inter.	[0.246,0.364]	[0.110,0.212]	[0.190,0.307]	[0.034,0.127]	[0.138,0.220]	[0.003,0.051]

Issue	big business (thermometer)	union (thermometer)	government standard of living
$\mu_{1992-2008}$	0.235	0.494	0.270
conf. inter.	[0.176,0.288]	[0.444,0.546]	[0.224,0.319]

Table 1: Estimation of Paramters; 95 percent confidence interval

Table 1 reports the values and 95 percent confidence interval (obtained by using bootstrap resampling) of λ and μ. All coefficients are significant on the 95 percent level, and the direction in which issue preferences translate into cultural and economic positions is always as expected: A voter is more economically conservative (i.e., low θ) if he likes big business; dislikes unions; and feels that government should not provide guaranteed jobs. He is more culturally conservative (i.e., high δ) if he likes the military; dislikes 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.

In terms of their weight for the determination of the economic index, 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); the questions of race and affirmative action (about 40 percent) and the women-specific questions (about 25 percent).[17] Note that weekly church attendance,

[17]The reader may wonder about the weight of the seemingly quaint and today mostly uncontroversial "role-of-women" question for the determination of social conservatism. The reason is that, exactly because an equal rights role of women is uncontroversial with most voters, a more conservative opinion on this issue has become a really strong signal for a respondent's cultural position.

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.

6.2 Platform Differentiation

To analyze changes in platform divergence, recall from equation (9) in Section 4.2 that the model identifies changes in the parties' policy distance relative to the corresponding 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 3 displays the results for cultural and economic positions.

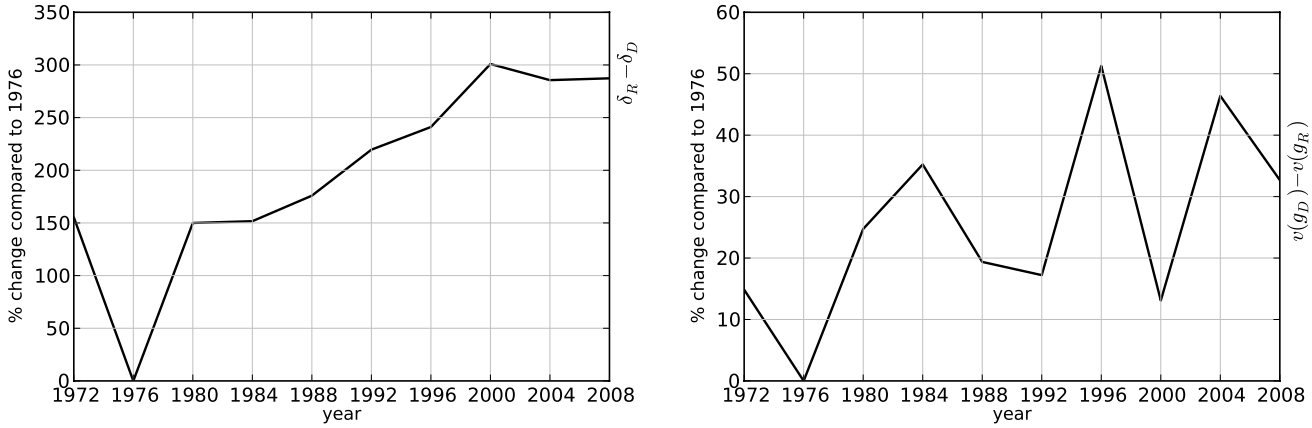

Figure 3: Cultural and economic policy divergence of candidates, 1972 to 2008

The distance between the two parties' cultural positions, $\delta_R - \delta_D$, relative to 1976, 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 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 effect of policy divergence on voter behavior. Figure 4 displays the values of δ and θ for all voters, together with the voter's choice (gray for Republican, black for Democrat). The left panel is for the 1976 election, the right one for the 2004 election. In both panels, the separating line divides voters who are more likely to vote for the Republican (below the line) from those more likely to vote for the Democrat (above the line), with types on the line having an implied probability of voting Republican or Democrat that is exactly $1/2$.

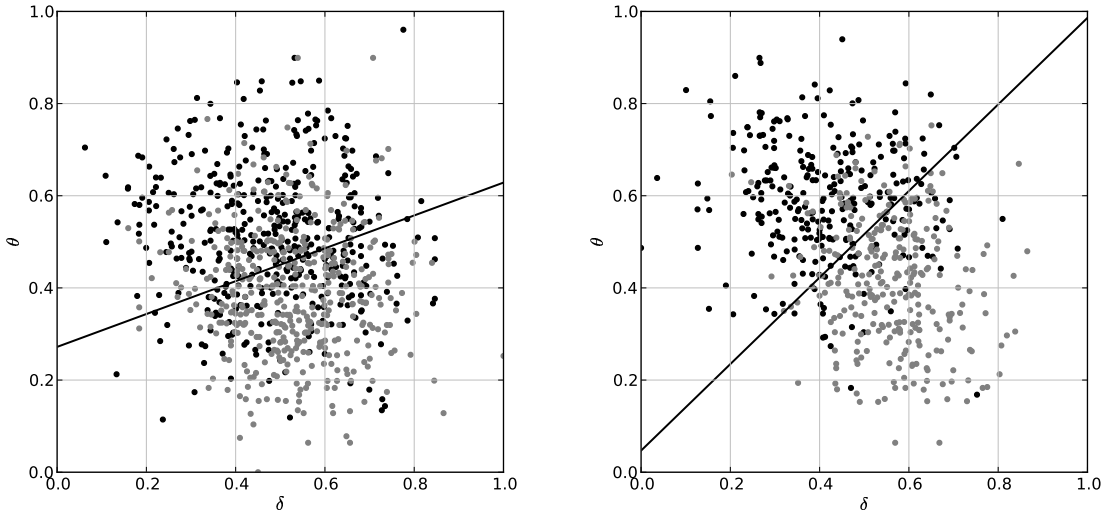

Figure 4: Voter preferences and vote choices in the 1976 (left) and 2004 (right) U.S. Presidential elections. Democratic voters in black, Republican ones in blue

Two features are evident from Figure 4. First, the ideological separation between Democrats and Republicans is much sharper in 2004 than in 1976. Clearly, this follows from policy divergence, both on economic and on cultural issues, being substantially stronger in 2004. We elaborate on this finding in Section 6.3.

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, so that social liberals primarily vote 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.

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.

Figure 5 displays the development of the slope k. After the initial decrease from 1972 to 1976, the relative importance of cultural issues starts to increase and reaches a high point in 2000, remaining relatively high afterwards. The confidence intervals in Figure 5 indicate that, while election-to-election changes are often not statistically significant, the long-term trend definitely is.

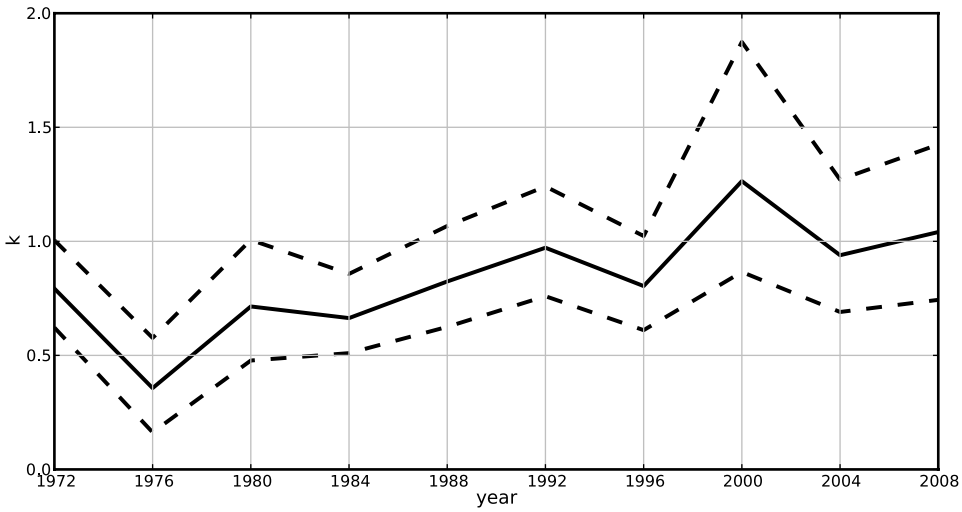

Figure 5: The development of k from 1972 to 2008, with 95% confidence intervals

Our results fit the narrative that Ronald Reagan's success as a conservative in 1980 against Carter was a key turning point in American politics that initiated a process of ideological realignment of the parties. After the relatively unpolarized 1976 election, cultural policy divergence in 1980 rebounds to the 1972 level and climbs steadily until plateauing out in 2000.

It is interesting to note that this sorting of conservatives and liberals into the two parties starts with Reagan's success in 1980, but is a long process rather than a one-time shock, as evidenced by the time series of k. 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.[18] 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.

6.3 Polarization, Radicalization and Sorting of the Electorate

We now return to the observation in the previous subsection that the increased policy divergence implies that voters' policy preferences become a better predictor of their voting behavior. As proposed in Section 4.4, 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.

The left panel of Figure 6 shows the development of Ψ over the last 10 presidential elections, and the parallels to cultural policy divergence in Figure 3 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' basic cultural and economic 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.

The right panel of Figure 6 shows how much of the total prediction success could be achieved if we knew only a voter's answers to the economic questions or to the cultural questions, respectively, expressed as a percentage of Ψ. So, for example, in 2008, knowing only the answers to the economic questions would result in a $\Psi_{2008,\text{ econ only}}$ that is about 79 percent of the size of Ψ_{2008}; knowing only the answers to the cultural questions would result in a Ψ that is about 87 percent of the size of Ψ_{2008}. Clearly, this increase in "cultural Ψ" reflects the increase in k, due to stronger policy differences on cultural issues.

Interestingly, in the first four elections, the economic questions alone explain much more of the total polarization measure than the cultural questions, and around 90 percent of the overall size

[18] See http://en.wikipedia.org/wiki/Party_switching_in_the_United_States.

Figure 6: Left: Polarization from 1972 to 2008, with 95% confidence intervals ; Right: Percentage of polarization explained by only economic (dashed) and only cultural (solid) preferences

of Ψ. In contrast, in the last five elections, both economic and cultural issues alone each account for around 80 percent of the total. In this sense, we can say that economic and cultural issues (as measured by the NES) are of roughly equal importance in determining a voter's vote choice.

It is instructive to compare the development of polarization in the left panel of Figure 6 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).[19] Overall, this external validation confirms that Ψ measures what has been interpreted as mass polarization in the existing literature.

The main advantage of Ψ relative to these existing measures is, though, that we can decompose the change in Ψ into the effects due to sorting and radicalization. Sorting $\Delta S(t)$ (defined in Section 4.3) isolates the effect of changes in platforms, holding fixed the distribution of political

[19]The 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.

preferences in society at the level of the previous election. Radicalization $\Delta R(t)$ isolates the effect of a changed voter preference distribution, holding fixed the candidates' platforms.

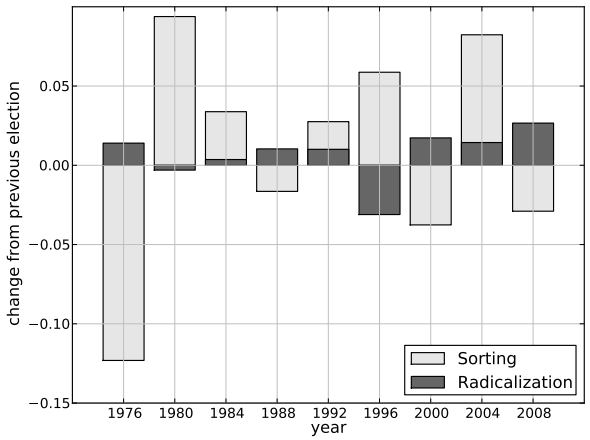

Figure 7: Sorting and radicalization contributions to polarization, 1972-2008

Figure 7 plots $\Delta S(t)$ and $\Delta R(t)$. 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 equals $\Delta\Psi_t$. In the other years, we draw both radicalization and sorting starting from zero, and $\Delta\Psi_t$ is equal to the difference between the positive and the negative column.

Clearly, 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, and candidates change from election to election, while the electorate remains mostly the same as in the previous election.

Interestingly, while parties became a lot more differentiated throughout the 1980s and 1990s so that sorting increased substantially, there is very little overall radicalization: The aggregate radicalization effect between 1976 and 1996 in Figure 7 is very close to zero. Thus, the conservative revolution affecting the political elite had arguably very little effect on the preference distribution of the American electorate at large. This seems to have changed with more substantial increases in radicalization in the last three elections, which 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 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.

Our measure of radicalization $\Delta R(t)$ captures changes in the voter preference distribution. 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. Obviously, increases in both the standard deviation of δ and of θ translate into positive $\Delta R(t)$, but 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. This confirms the results of DiMaggio et al. (1996), Fiorina et al. (2006) and Fiorina and Abrams (2008) who all find that overall issue preferences of American voters have remained mostly stable over time.

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

However, the correlation between economic and cultural conservatism among voters has increased from a low of 0.18 in 1976 to 0.46 in 2008, and the increased correlation between δ and θ is 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;[20] 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).

7 The ideological preferences of non-voters: What if everybody voted?

In most democracies, voting is a voluntary activity, and many citizens choose not to vote. How do the ideological preferences of non-voters look like, and what are the partisan consequences of abstention? Because legislatures can make it easier or harder to vote (e.g., automatic registration, "motor voter" laws or mail-in voting on the one side, voter-ID laws on the other side), these are not only intellectually interesting questions, but the answers have important policy consequences.

The theoretical literature that analyzes the desirability of encouraging citizens to vote typically focuses on a setting where there are no partisan differences in the distribution of voting costs and varies assumptions about the partisan composition and information status of the electorate (Börgers, 2004; Krasa and Polborn, 2009; Ghosal and Lockwood, 2009; Krishna and Morgan, 2012). Encouraging voting in these models may have positive or negative welfare effects, but there are no partisan benefits of increased turnout rates.

In practice, the conventional wisdom among journalists and political pundits is that, because non-voters in the U.S. belong disproportionately often to ethnic minorities and economically disadvantaged strata – groups that support Democrats by a substantial margin – an increase in turnout

[20] We do not have a formal test of what is driving the increase in correlation between δ and θ, but it is interesting to speculate whether it is related to partisan news media and talk radio. Maybe, voters learn from the internally consistent world view that Fox News and MSNBC provide that cultural conservatives and cultural liberal "should" also be economically conservative and economically liberal, respectively.

would be beneficial for Democrats. A revealed preference argument suggests that this belief is shared among political practitioners: While laws facilitating voting are usually passed by legislatures controlled by the Democrats, laws making voting more difficult are usually passed by legislatures controlled by Republicans.

Surprisingly, quantitative research in political science suggests that the impact of increased turnout on which candidate wins in Senate elections or Presidential elections is minimal (DeNardo, 1980; Tucker et al., 1986; Citrin et al., 2003; Sides et al., 2008). For example, Citrin et al. (2003) estimate, for 91 U.S. Senate elections in the 1990s, that the Democratic vote share would only have increased by 0.7 percent from 48.4 percent to 49.1 percent if *all* registered voters had voted. Their analysis is based only on demographic data of voters from exit polls (such as gender, race and income), and assumes that non-voters who share these demographic characteristics would vote for the parties at the same rate as the corresponding exit poll voters.

These empirical results create a substantial puzzle: Since any practical law that makes voting more difficult will not lead to dramatic changes in the overall participation rate, the practical importance of such laws would appear to be extremely small and not worth spending any effort on promoting them. This is especially true since laws that make voting more difficult also affect current voters and are likely unpopular with them because they increase their cost of voting.

In contrast to the papers cited above, we analyze how the preference distribution of non-voters interviewed in the National Election Survey differs from that of voters, and how these non-voters would have voted (probabilistically) if they voted according to the same model as their ideological compatriots who voted.

The left panel of Figure 8 displays the Democratic share of the two-party vote in the electorate at-large (the dotted line) and among NES respondents who voted; the solid line is derived from a raw count of the respondents' voting decision, and the (essentially coinciding) dashed line is derived by predicting the behavior of all NES voters as implied by their (δ, θ) position.[21] Note that the NES sample relatively closely reflects the actual election outcomes, except in 2008.

[21] The main point of this comparison is to show that imputing voting decisions from ideological positions of voters leads, on aggregate, to predicted vote shares very close to the actual ones. This is important because we do not observe the "actual voting decisions" that non-voters would make, just their ideological preferences.

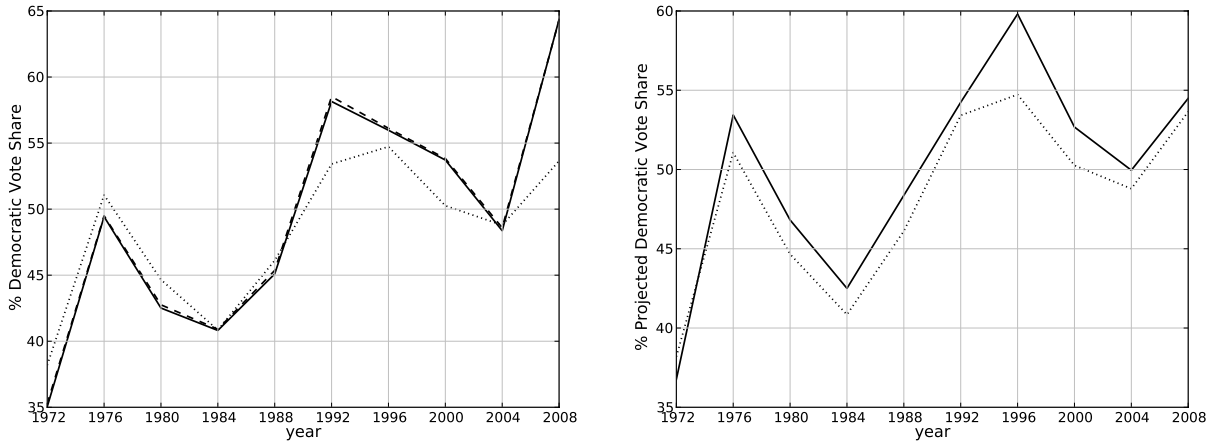

Figure 8: Share of Democrats among voters (left) and if all registered voters voted

In the right panel, the dotted line is again the actual election outcome in the electorate at-large, and the solid line shows the election outcome if all eligible voters would actually have voted. To calculate this prediction, we proceed as follows: First, we calculate the implied probability of voting for the Democratic candidate among voters and non-voters. From this, we calculate the percentage of "excess Democrats" among NES non-voters. For example, if 49 percent of voters and 58 percent of non-voters in the NES are predicted to vote for the Democrat, there are 9 percent excess Democrats among non-voters. We then calculate a predicted Democratic share among non-voters as equal to the Democratic share in the actual election, plus the "excess Democrats" percentage from the comparison of NES voters and non-voters. Thus, if the Democratic share in the actual election results was only 47 percent (rather than the 49 percent in the NES sample), then the predicted Democratic share among non-voters is $47 + 9 = 56$ percent. Finally, we calculate a weighted average of the Democratic percentage in the actual election results and the predicted Democratic share among non-voters, where the weights are based on the actual turnout rates taken from `http://www.presidency.ucsb.edu/`. For example, if the turnout rate was $2/3$, then the predicted Democratic share if all voters voted is $(2/3) \times 47 + (1/3) \times 56 = 50$ percent.

Since 1976, Democrats would have performed on average about two to three percentage points better, if all voters had participated. This gap is largest in 1996 and 2000, and would have changed the election outcome in 2000 and possibly in 2004. The narrowing of the gap in 2004 and 2008

can be interpreted as a result of improved Democratic turnout operations in these years, essentially already tapping a large part of their potential voter pool.

Thus, our findings here support the intuitive view that Democrats would benefit from increased turnout, and this effect is considerably stronger than the one found in the papers cited above. Intuitively, the reason is that the extent of the difference that a study finds depends on two factors related to the characteristics on which the study conditions: First, how good are these characteristics in predicting voting behavior, and second, how different are the composition of the two groups of voters and non-voters with respect to these characteristics? Apparently, the demographic characteristics used in the studies above are relatively poor predictors of voting behavior, and this leads to an underestimate of the partisan effects of an increased turnout rate.

8 Robustness

8.1 Overview

In the following we discuss four different robustness issues. We start with providing a summary of the detailed analysis that can be found below.

First, in our analysis, we assume that the standard deviation of ξ does not change over time.[22] In a probit 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 – if we write the minimization problem of a probit regression, but assume that the probability of voting Republican is $\Phi_\sigma(\alpha+\beta x)$ (where Φ_σ is the cdf of a $N(0,\sigma)$ distributed random variable), then the objective function is homogeneous of degree zero in (α,β,σ). Thus, σ is not determined and can be normalized to one, without loss of generality.

In contrast, when we interpret the change of regression coefficients (or functions of regression coefficients) over time, we effectively assume that the standard deviation of idiosyncratic prefer-

[22] We do not need to make any assumption about the average value of ξ in the population, i.e. the average net valence of candidates is allowed to vary over time.

ence shocks is constant over time. This is a standard assumption in methods that compare regression coefficients over time and is usually not even discussed. For example, in their discussion of the DW-Nominate method, Poole and Rosenthal (2011), p. 27, note in passing that "We assume the signal-to-noise ratio [their expression for the error term] is constant across all of American history".[23] However, we can use information from personal like/dislike questions from the NES to normalize σ_ξ to a non-constant time series that may better reflect changes in the distribution of idiosyncratic personal preferences.

Second, we analyze what happens if voters' true economic and cultural positions do not just depend on the positions on those questions that we have data on, but also on other issues. We show that such a misspecification would not bias the estimation of k. Furthermore, the estimate of elite polarization would be biased downward, implying that our result of substantial elite polarization would be strengthened further.

Third, our measure of "cultural issues" lumps together all available non-economic policy questions in the NES. This has the interpretative advantage of providing for just one marginal rate of substitution between economic issues and all other issues, but may be problematic if policy divergence develops unevenly in different cultural policy areas. Therefore, we analyze the robustness of our results to the aggregation of different cultural issues by treating all cultural questions as separate issues, so that the weights of these issues can change freely between elections.

Finally, we compare our estimates of policy divergence with the naive measure obtained from the NES question that asks respondents to place presidential candidates on a left-right-spectrum. Apart from the fundamental problem discussed earlier that many respondents have difficulty placing candidates on a left-right spectrum, we show that different voters disagree considerably about the position of candidates, and that the naive measure cannot capture the historical developments.

[23] Bartels (2006) takes a similar approach.

8.2 Changes in the Variance of Valence ξ

If the standard deviation of idiosyncratic preference shocks is constant over time, we can interpret our empirical results as evidence of policy divergence. If, instead, one allows for σ_ξ to vary over time, the interpretation of the policy divergence results can change; for example, if one were to assume that σ_ξ decreased considerably over time (i.e., the size of the average idiosyncratic preference shock decreased), then one would have to think of overall policy divergence between parties as relatively constant (though there still would have to be an increase of cultural divergence relative to economic divergence). If, instead, σ_ξ increases over time, the divergence effects would be magnified relative to the basic model. The mathematical logic behind our model (and, more generally, intertemporal probit models) does not allow us to isolate one of these interpretations as the "true" one any more than a relativist physicist can determine an absolute coordinate system.[24]

This said, what is a natural way of thinking about the temporal development of σ_ξ in our context? The net-valence term ξ is determined by the voters' interpretation of candidate traits that are not directly linked to the candidate's economic or cultural platform, and the NES contains several question about such characteristics that go back sufficiently many years to enable a comparison across different elections: VCF0354 – VCF0356 and VCF0366 – VCF0366 ask, respectively, whether the Democrat and Republican presidential candidates are knowledgeable, moral and provide strong leadership. Each of these variables is measured on a 4-point scale, and if we denote the responses of voter j to the questions about the Democratic and Republican candidate at time t by $X^j_{i,t}$, $Y^j_{i,t}$, $i = 1, 2, 3$, respectively, then $Z^j_t = \sum_{i=1}^3 (X^j_{i,t} - Y^j_{i,t})$ is a useful proxy that is proportional to the net valence of the Democratic candidate that voter j perceives. We can then compute the standard deviations $\sigma(Z_t) = \sqrt{E[(Z^j_t - E(Z^j_t))^2]}$ for the presidential election years from 1980 (the first year for which these data are available) to 2008, which gives the following values: 3.10, 3.00, 2.62, 3.08 3.13 3.20, 4.21, and 4.05.

The solid line in Figure 9 recalculates the time series from Figure 3 in Section 6, using these

[24] I.e., if the physicist pushes the gas pedal in a car, does the car accelerate in the direction it is pointing, or does the car stand still, but the trees move faster in the opposite direction? Modern physics is built on the notion that there is no absolute coordinate system, so we cannot say which of the two statements is in any absolute sense "true", but it is still the case that certain interpretations are more natural than others in certain applications.

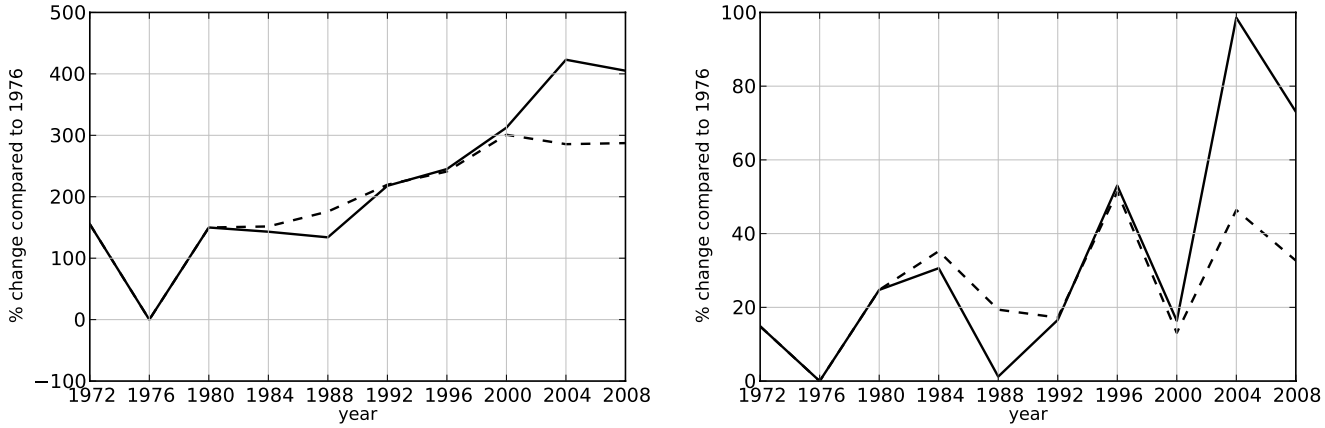

Figure 9: Cultural and economic policy divergence of candidates, 1972 to 2008, when σ_ξ changes. Dashed line: constant σ_ξ.

standard deviations for σ_ξ. For comparison, we plot the values derived from assuming that $\sigma(\xi)$ is constant (i.e., the values of Section 6) as a dashed line. Note that the two curves are very close to each other until 2000, and thus, the overall picture of the development until this time is qualitatively unchanged. However, for 2004 and 2008, the adjusted curve displays even more policy divergence than in the basic model where σ_ξ is assumed to be constant.

8.3 Missing Policy Questions

In the main text above, we assume that questions X_i and Y_i contain all relevant policy information and can therefore be mapped perfectly into policy positions δ and θ. Now suppose that there exist questions $\hat{X}_i$ and $\hat{Y}_i$ that also determine a voters position on cultural and economic issues, but are not included in the NES survey. Suppose, the answers to these missing questions can be decomposed into a combination of answers to the existing questions plus an independent term that is normally distributed. Then δ and θ are given by $\delta = \sum_{i=1}^{n} \lambda_i X_i + \epsilon_X$ and $\theta = \sum_{i=1}^{m} \mu_i Y_i + \epsilon_Y$. Random variables ϵ_X and ϵ_Y are mutually independent, and also independent from the X_i and Y_i.

Let $\tilde{\varepsilon} = \epsilon + k\epsilon_X + \epsilon_Y$. Then $\tilde{\varepsilon}$ is normally distributed, with mean zero and standard deviation

$\tilde{\sigma} = \sqrt{\sigma + k\sigma_X + \sigma_Y}$. A citizen votes Republican if

$$\sum_{i=1}^{m} \mu_i Y_i - k \sum_{i=1}^{n} \lambda_i X_i - a - \tilde{\varepsilon} < 0, \tag{20}$$

which is identical to equation (6) if we replace $\tilde{\varepsilon}$ by ε.

Similarly, if we replace ϵ_t by $\tilde{\epsilon}_t$ and σ_t by $\tilde{\sigma}_t$ in Theorem 1, it is clear that only the third statement is affected, i.e., the formula in the Theorem now provides the standard deviation $\tilde{\sigma}_t$ instead of σ_t. Most importantly, the estimation of k is completely unaffected by missing questions.

In our above analysis the missing questions matter when we investigate changes in elite polarization. In particular, (9) is replaced by

$$\delta_D - \delta_R = \frac{\sigma_\xi k}{2\sqrt{\tilde{\sigma}^2 - k\sigma_X^2 - \sigma_Y^2}}, \text{ and } v(g_D) - v(g_R) = \frac{\sigma_\xi}{\sqrt{\tilde{\sigma}^2 - k\sigma_X^2 - \sigma_Y^2}} \tag{21}$$

Suppose that σ_X and σ_Y have remained constant, i.e., the problem of missing questions has not changed. Since k has increased over the sample years, this decreases the denominator for both expressions and hence raises both $\delta_D - \delta_R$ and $v(g_D) - v(g_R)$. Similarly, $\tilde{\sigma}$ decreases over the sample years, and hence (21) increases at a faster rate than (9). Thus, both effects reinforce the difference between the candidates' positions, resulting in a larger increase of elite polarization than in Figure 3. In other words, missing questions would strengthen our results on elite polarization.

8.4 Separating the "cultural" issues

Our measure of "cultural issues" lumps together all "non-economic" policy questions that are continuously available 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 4, 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, ag-

Figure 10: Polarization from 1972 to 2008 when all cultural questions are separate issues (solid line; previous result is dashed line)

gregating all non-economic policies into one dimension may be problematic if 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 10 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 polarization are also almost the same as in the basic model, showing the robustness of these results.

8.5 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 may misunderstand the question about a position in an abstract horizontal policy space (e.g., what positions are really "moderately conservative"?). In contrast, the concrete policy questions in the NES 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 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.

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

There is also no statistically significant time trend of the average Democratic or Republican position. 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 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 perceived divergence of platforms, however, it may also mean that more voters became partisans and place the candidates further apart on the left-right spectrum. Again, absent a model we cannot get clear answers from the data.

9 Discussion and concluding remarks

9.1 Existing methods of position measurement

It is useful to contrast our method of position measurement in presidential elections with methods in the existing literature. As already explained in the introduction, there are two main differences: First, while existing methods infer a legislator's positions from his voting behavior, our measure is based on a "revealed preference approach," specifically on how voters *perceive* the candidates' positions. Second, our method does not only provide position estimates for candidates, but also measures for the relative importance of economic and cultural issues, and allows us to simultaneously analyze both elite and mass polarization in the same framework.

We now discuss in some more detail Poole and Rosenthal's DW-nominate method which measures legislators' positions, based on their votes in Congress. Legislators are assumed to have utility functions of the form $\exp(-d^2) + \varepsilon$, where d is the distance between their respective ideal points and the proposed policy, and ε is a noise term. Poole and Rosenthal estimate positions using

role call data for one and two dimensions, show that for many years adding a second dimensions increases the fit of their model by about 10% (see Average Proportional Reduction in Error in Table 3.1 of Poole and Rosenthal (2011)) and argue that this implies that policy divisions in Congress are almost one-dimensional.

Note that, in their model, legislative votes are not "assigned" to be about economic or cultural issues. Intuitively, their approach corresponds to selecting a coordinate system to best fit their data, where the "first" dimension combines votes on many issues in a way that captures "most" policy differences, and the "second" dimension is orthogonal to the first one. Clearly, their implicit coordinate system may change from one Congress to the next, which makes distance comparisons across different years more challenging, and does not show which type of issues have become more important for polarization in Congress.

Like Poole and Rosenthal, we can only get a distance between candidates, and the distance can only be determined up to a scaling factor. Intuitively, this is equivalent to having a list of distances between different cities where the distance measure itself is not specified (e.g., it could be km or miles). In order to get an actual map, one would have to choose some normalization, e.g., on a one-dimensional map the location of two cities. This is precisely what Poole and Rosenthal do by assigning the most liberal politician to -1 and the most conservative one to $+1$.

However, this means that comparing the average DW-nominate scores of Democrats and Republicans in Congress across time generates an interpretative challenge: When the average Republican's estimated position in Congress increases over time, is this because their policy positions in some true (unnormalized) policy space become on average more extreme, or is it maybe that the most extreme Republicans become more moderate over time in this true policy space, relative to the bulk of their caucus? In both cases, the average Republican DW-Nominate score would increase.[25] In contrast, we measure voters' economic and cultural positions always based on the same set of fundamental questions, so voter preference distributions are intertemporally comparable. Furthermore, our policy divergence estimates are based on the behavior of these voters.

[25]To adjust for this issue, some authors use additional information such as ADA scores to make intertemporal adjustments (cf., Groseclose, Levitt and Snyder (1999)).

Also, note that the DW-Nominate method assumes that the size of the error term is constant throughout time. Consider the finding that the "positions" of the median Democrat and the median Republican in Congress have moved away from each other since the 1970s. This is generally interpreted to be a meaningful statement about the parties' "political positions" having become more polarized, but this interpretation also relies on the assumption of a constant error term. Without the assumption of a constant error term, all one can interpret from the DW-Nominate results is that in votes that split Congress 50/50, the vote of the median Democrat and of the median Republican are less likely to coincide today than in the 1970s. In principle, such a finding could also arise if the (again, meaningfully defined) "positions" of the median Democrat and the median Republican remained constant since the 1970s, but the standard deviation of the error term has decreased. To be sure, there are good substantive reasons for any careful observer of politics to believe that there was meaningful polarization in Congress rather than a spurious decrease in the error term, but in the absence of any fixed yardstick (e.g., a measure of voters' willingness to pay for their preferred outcome in each vote), the results of any position measurement method – ours as well as DW-Nominate – indicate relative importance of the measured position to all other factors. It is important to keep this in mind when interpreting the results, but, in our opinion, should not lead us to throw out the baby with the bath water and dispense with all position measurements in politics.

Using the DW-Nominate method on Presidential candidates directly is difficult because the candidates rarely serve in the same legislature before running for President (Obama vs. McCain in 2008 was the only exception to this in the recent past). There is a variety of techniques that attempt to compare candidates across different institutions. Many of them relying on"bridge actors" that have served in multiple institution, thus providing a link between the candidates. For executives, such as the president, statements about legislations have been used to compare their ideological position to those of other legislators, or use them as "bridges" to evaluate the ideological position of actors in different institutions (cf., Bailey and Chang (2001)).

An important distinction between our model and existing methods is that the weight of different issues in our approach is determined endogenously by how much they influence the voters' choice, while other methods usually impose weights exogenously. For example, interest group measures

of legislator positions (such as ADA-scores) select a certain number of legislative votes that receive equal weights, and all other votes receive a zero weight. In the DW-Nominate method, all legislative votes implicitly have the same weight for the determination of positions as no adjustment for the importance of a vote is made. As a consequence, if there were two policy dimensions that were equally important in an absolute sense, but Congress was voting *more often* on issues related to the first dimension than on those related to the second dimension, the DW-Nominate method would classify the first dimension as more important (as it explains more Congressional votes).

9.2 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 show that the increase in policy divergence precedes any significant radicalization of the electorate by about 20 years. Thus, models in which the electoral preference distribution play the determinant role for where candidates choose their position, such as entry deterrence models, cannot account for the observed change in policy divergence. Of course, these models still provide valuable intuition for some potential reasons of policy divergence.

In policy-motivation models, divergence arises from the interplay of policy motivated candidates and uncertainty about the median voter's preferences. 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.[26] In both parties the number of states that hold primaries increases throughout the 1970s and 1980s, increasing the power of regular party members at the expense of party elites.[27] It is likely that party elites are more interested in winning *per se* (as a winning 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 over time: As long as both parties' leaders select moderate candidates, the incentive for regular voters to choose their party membership based on their ideology is limited. 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.

9.3 Concluding remarks

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 distinct, then most citizens vote primarily according to their policy preferences, whereas, if candidate positions are close, then they choose candidates primarily based on non-policy attributes. 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.

[26] See `http://en.wikipedia.org/wiki/Democratic_National_Convention`.

[27] For example, just 28 states had primaries or caucuses in 1976, while in 1996, 47 states did.

Our model allows us to measure the development of policy divergence between Democratic and Republican Presidential candidates, both on economic and on cultural issues. We find that, since Ronald Reagan's victory in 1980, the two parties have diverged substantially, in particular on cultural issues.

We also use the model to define a concept of mass (voter) polarization. The intensity with which voters care about elections and the extent to which their voting decisions depend on their policy preferences depend positively on the policy differences between the competing parties, and on how extreme voters' policy preferences are. We can thus identify how much party policy divergence and voter radicalization contribute to changes in overall mass polarization.

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 methods that we have developed in this paper will prove useful in this long-term project.

10 Appendix

10.1 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,

$$\sum_{i=1}^n \tilde{\lambda}_i\tilde{X}_i = \sum_{i=1}^n \lambda_i X_i \sum_{i=1}^n |\tilde{\lambda}_i| + \sum_{i=1}^n \min\{\tilde{\lambda}_i, 0\}, \text{ and } \sum_{i=1}^m \tilde{\mu}_i\tilde{Y}_i = \sum_{i=1}^m \mu_i Y_i \sum_{i=1}^m |\tilde{\mu}_i| + \sum_{i=1}^m \min\{\tilde{\mu}_i, 0\} \quad (22)$$

Since $\delta = \sum_{i=1}^n \lambda_i X_i$, and $\theta = \sum_{i=1}^m \mu_i Y_i$ equation (22) immediately implies (15).

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

Note that (19) and (17) imply

$$\frac{a_t}{\sigma_t} = \tilde{a}_t - (1 + \rho_t) \sum_{i=1}^m \min\{\tilde{\mu}_i, 0\} + (1 + \alpha_t) \sum_{i=1}^n \min\{\tilde{\lambda}_i, 0\}. \quad (23)$$

(22) implies

$$(1 + \alpha_t) \sum_{i=1}^n \tilde{\lambda}_i\tilde{X}_i = (1 + \alpha_t) \sum_{i=1}^n \lambda_i X_i \sum_{i=1}^n |\tilde{\lambda}_i| + (1 + \alpha_t) \sum_{i \in N_\Lambda} \tilde{\lambda}_i = \frac{k_t}{\sigma_t} \sum_{i=1}^n \lambda_i X_i + (1 + \alpha_t) \sum_{i=1}^n \min\{\tilde{\lambda}_i, 0\} \quad (24)$$

$$(1 + \rho_t) \sum_{i=1}^m \tilde{\mu}_i\tilde{Y}_i = (1 + \rho_t) \sum_{i=1}^n \mu_i Y_i \sum_{i=1}^m |\tilde{\mu}_i| + (1 + \rho_t) \sum_{i \in N_M} \tilde{\mu}_i = \frac{1}{\sigma_t} \sum_{i=1}^n \mu_i Y_i + (1 + \rho_t) \sum_{i=1}^m \min\{\tilde{\mu}_i, 0\}. \quad (25)$$

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, (24), and (25) imply

$$\begin{aligned}
&\left(1 + \sum_{t=2}^{s} \alpha_t D_t\right) \sum_{i=1}^{n} \tilde{\lambda}_i \tilde{X}_i - \left(1 + \sum_{t=2}^{s} \rho_t D_t\right) \sum_{i=1}^{m} \tilde{\mu}_i \tilde{Y}_i + \sum_{t=1}^{s} \tilde{a}_t D_t \\
&\quad = \sum_{t=1}^{s} D_t (1 + \alpha_t) \sum_{i=1}^{n} \tilde{\lambda}_i \tilde{X}_i - \sum_{t=1}^{s} D_t (1 + \rho_t) \sum_{i=1}^{m} \tilde{\mu}_i \tilde{Y}_i + \sum_{t=1}^{s} \tilde{a}_t D_t \\
&\quad = \sum_{t=1}^{s} D_t \frac{k_t}{\sigma_t} \sum_{i=1}^{n} \lambda_i X_i - \sum_{t=1}^{s} \frac{D_t}{\sigma_t} \sum_{i=1}^{m} \mu_i Y_i \\
&\qquad\qquad + \sum_{t=1}^{s} D_t \left[\tilde{a}_t + (1 + \alpha_t) \sum_{i=1}^{n} \min\{\tilde{\lambda}_i, 0\} - (1 + \rho_t) \sum_{i=1}^{m} \min\{\tilde{\mu}_i, 0\}\right] \\
&\quad = \sum_{t=1}^{s} \frac{D_t}{\sigma_t} \left[\sum_{t=1}^{s} k_t D_t \sum_{i=1}^{n} \lambda_i X_i - \sum_{i=1}^{m} \mu_i Y_i + \sum_{t=1}^{s} D_t a_t\right],
\end{aligned}$$

where the last equality follows from (23). The two models are therefore equivalent. ■

10.2 The General Model

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

$$u(\delta, \theta, \xi_P) = \theta v(g_P) - c_P - \sum_{j=1}^{J} w_j (\delta_j - \delta_{j,P})^2 + \xi_P, \tag{26}$$

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

It is easy to verify that (3) generalizes to

$$\theta(\delta, \xi, g_D, g_R) = \frac{2 \sum_{j=1}^{J} \delta_j w_j (\delta_{j,R} - \delta_{j,D}) + c_D - c_R - \sum_{j=1}^{J} w_j (\delta_{j,R}^2 - \delta_{j,D}^2) + \xi}{v(g_D) - v(g_R)}. \tag{27}$$

Let

$$k_j = \frac{2 w_j (\delta_{j,R} - \delta_{j,D})}{v(g_D) - v(g_R)}, \quad a = \frac{c_D(g_D) - c_R(g_R) - \sum_{j=1}^{J} w_j (\delta_R^2 - \delta_D^2) + \bar{\xi}}{v(g_D) - v(g_R)}, \tag{28}$$

where $\bar{\xi} = E[\xi]$. Let $k = (k_j)_{i=1,\ldots,J}$ Then (27) implies that the separating hyperplane is given by $\theta = k \cdot \delta + a$, and a citizen votes Republican if and only if

$$\theta - k \cdot \delta - a - \varepsilon < 0. \tag{29}$$

Let $X_{i,j}$, $i = 1, \ldots, n_j$, $j = 1, \ldots J$ and Y_i, $i = 1, \ldots, m$ be random variables that describe the answers to survey questions on issues $j = 1, \ldots, J$ 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 $j = 1, \ldots, J$ and $\sum_{i=1}^{m} \mu_i = 1$.

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

$$\Phi\left(\sum_{t=1}^{s} \frac{D_t}{\sigma_t}\left[\sum_{j=1}^{J}\left(\sum_{i=1}^{s} D_t k_{j,t}\right)\left(\sum_{i=1}^{n} \lambda_{i,j} X_{i,j}\right) - \sum_{i=1}^{m} \mu_i Y_i + \sum_{t=1}^{s} D_t a_t\right]\right). \tag{30}$$

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

$$\Phi\left(\sum_{j=1}^{J}\left(1 + \sum_{t=2}^{s} \alpha_{j,t} D_t\right)\left(\sum_{i=1}^{n} \tilde{\lambda}_{i,j} \tilde{X}_{i,j}\right) - \left(1 + \sum_{t=2}^{s} \rho_t D_t\right)\left(\sum_{i=1}^{m} \tilde{\mu}_i \tilde{Y}_i\right) + \sum_{t=1}^{s} \tilde{a}_t D_t\right), \tag{31}$$

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

$$z_\ell = \sum_{\ell=1}^{L}\sum_{j=1}^{J}\left(1 + \sum_{t=2}^{s} \alpha_{j,t} d_{t,\ell}\right)\left(\sum_{i=1}^{n} \tilde{\lambda}_i \tilde{x}_{i,j,\ell}\right) - \left(1 + \sum_{t=2}^{s} \rho_t d_{t,\ell}\right)\left(\sum_{i=1}^{m} \tilde{\mu}_i \tilde{y}_{i,}\right) + \sum_{t=1}^{s} \tilde{a}_t d_{t,j}, \tag{32}$$

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

$$\max_{\{\alpha_i,\rho_i|i=2,\ldots,s\},\{\tilde{a}_i|i=1,\ldots,s\},\{\tilde{\lambda}_{i,j}|i=1,\ldots,n,\,j=1,\ldots,J\},\{\tilde{\mu}_i|i=1,\ldots,m\}} \sum_{j=1}^{J} v_j \ln \Phi(z_j) + (1-v_j)\ln\Big(1-\Phi(z_j)\Big). \qquad (33)$$

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 $\tilde{\lambda}_{i,j}$.

References

Abramowitz, Alan and Kyle Saunders, "Ideological Realignment in the U.S. Electorate," *Journal of Politics*, 1998, *60* (03), 634–652.

_ **and** _ , "Is polarization a myth?," *Journal of Politics*, 2008, *70* (2), 542–55.

Ansolabehere, Stephen, Jonathan Rodden, and James M. Snyder, "Purple America," *Journal of Economic Perspectives*, 2006, *20* (2), 97–118.

Bailey, Michael and Kelly H. Chang, "Comparing presidents, senators, and justices: interinstitutional preference estimation," *Journal of Law, Economics, and Organization*, 2001, *17* (2), 477–506.

Bartels, Larry M., "Partisanship and voting behavior, 1952-1996," *American Journal of Political Science*, 2000, *44* (1), 35–50.

_ , "What's the matter with 'What's the matter with Kansas?'," *Quarterly Journal of Political Science*, 2006, *1* (2), 201–26.

_ , *Unequal democracy: The political economy of the new gilded age*, Princeton Univ Pr, 2010.

Bernhardt, Dan, John Duggan, and Francesco Squintani, "Private Polling in Elections and Voter Welfare," 2006. mimeo, University of Illinois.

Börgers, Tilman, "Costly voting," *American Economic Review*, 2004, pp. 57–66.

Callander, Steven, "Electoral Competition in Heterogeneous Districts," *Journal of Political Economy*, 2005, *113*, 1116–1145.

_ , "Political Motivations," *Review of Economic Studies*, 2008, *75*, 671–697.

Calvert, Randall L., "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence," *American Journal of Political Science*, 1985, *29*, 69–95.

Castanheira, Micael, "Why Vote For Losers?," *Journal of the European Economic Association*, 2003, *1* (5), 1207–1238.

Citrin, Jack, Eric Schickler, and John Sides, "What if everyone voted? Simulating the impact of increased turnout in Senate elections," *American Journal of Political Science*, 2003, *47* (1), 75–90.

Degan, Arianna, "Candidate valence: Evidence from consecutive presidential elections," *International Economic Review*, 2007, *48* (2), 457–482.

DeNardo, James, "Turnout and the Vote: The Joke's on the Democrats," *American Political Science Review*, 1980, pp. 406–420.

DiMaggio, Paul, John Evans, and Bethany Bryson, "Have Americans' Social Attitudes Become More Polarized?," *American Journal of Sociology*, 1996, *102*, 690–755.

Downs, Anthony, *An economic theory of democracy*, Harper, 1957.

Duclos, Jean-Yves, Joan Esteban, and Debraj Ray, "Polarization: concepts, measurement, estimation," *Econometrica*, 2004, *72* (6), 1737–1772.

Esteban, Joan and Debraj Ray, "Conflict and distribution," *Journal of Economic Theory*, 1999, *87* (2), 379–415.

_ and _ , "Linking conflict to inequality and polarization," *American Economic Review*, 2011, *101* (4), 1345–1374.

Esteban, Joan-Maria and Debraj Ray, "On the measurement of polarization," *Econometrica*, 1994, *62* (4), 819–851.

Feddersen, Timothy and Faruk Gul, "Polarization and Income Inequality: A Model of Unequal Democracy," 2013. mimeo, Northwestern University.

Fiorina, Morris P. and Samuel J. Abrams, "Political polarization in the American public," *Annual Review of Political Science*, 2008, *11* (1), 563–588.

_ , _ , and Jeremy C. Pope, *Culture war? The myth of a polarized America*, Pearson Longman, 2006.

Frank, Thomas, *What's the matter with Kansas?: How conservatives won the heart of America*, Holt Paperbacks, 2005.

Gelman, Andrew, Boris Shor, Joseph Bafumi, and David Park, "Rich state, poor state, red state, blue state: What's the matter with Connecticut?," *Quarterly Journal of Political Science*, 2008, *2*, 345–367.

Ghosal, Sayantan and Ben Lockwood, "Costly voting when both information and preferences differ: is turnout too high or too low?," *Social Choice and Welfare*, 2009, *33* (1), 25–50.

Greenberg, Stanley B., *The two Americas: Our current political deadlock and how to break it*, St. Martin's Griffin, 2005.

Groseclose, Timothy, Steven D. Levitt, and James M. Snyder, "Comparing interest group scores across time and chambers: Adjusted ADA scores for the US Congress," *American Political Science Review*, 1999, *93* (1), 33–50.

Gul, Faruk and Wolfgang Pesendorfer, "Partisan Politics and Aggregation Failure with Ignorant Voters," *Journal of Economic Theory*, 2009, *144*, 146–174.

Harbridge, Laurel and Neil Malhotra, "Electoral Incentives and Partisan Conflict in Congress: Evidence from Survey Experiments," *American Journal of Political Science*, 2011, *55* (3), 494–510.

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

Hunter, James D., *Culture wars: The struggle to define America*, Basic Books, 1992.

Krasa, Stefan and Mattias K. Polborn, "Is mandatory voting better than voluntary voting?," *Games and Economic Behavior*, 2009, *66* (1), 275–291.

_ and _ , "Competition between Specialized Candidates," *American Political Science Review*, 2010, *104* (4), 745–765.

_ and _ , "Social ideology and taxes in a differentiated candidates framework," *American Economic Review*, 2014, p. forthcoming.

Krishna, Vijay and John Morgan, "Voluntary voting: Costs and benefits," *Journal of Economic Theory*, 2012.

Levendusky, Matthew, *The partisan sort: How liberals became Democrats and conservatives became Republicans*, University of Chicago Press, 2009.

Martinelli, Cesar, "Elections with Privately Informed Parties and Voters," *Public Choice*, 2001, *108* (1), 147–167.

McCarty, Nolan, Keith T. Poole, and Howard Rosenthal, *Polarized America: The dance of ideology and unequal riches*, MIT Press, 2006.

Palfrey, Thomas R., "Spatial Equilibrium with Entry," *Review of Economic Studies*, 1984, *51*, 139–156.

Poole, Keith T. and Howard Rosenthal, "The polarization of American politics," *Journal of Politics*, 1984, *46* (04), 1061–1079.

_ and _ , "A spatial model for legislative roll call analysis," *American Journal of Political Science*, 1985, *29* (2), 357–384.

_ and _ , *Congress: A political-economic history of roll call voting*, Oxford University Press, 2000.

_ and _ , *Ideology and congress*, Transaction Publishers, 2011.

Shogan, Robert, *War without end: cultural conflict and the struggle for America's political future*, Basic Books, 2002.

Sides, John, Eric Schickler, and Jack Citrin, "If Everyone Had Voted, Would Bubba and Dubya Have Won?," *Presidential Studies Quarterly*, 2008, *38* (3), 521–539.

Soubeyran, Raphael, "Does a disadvantaged candidate choose an extremist position?," *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, 2009, *93/94*, 327–348.

Tucker, Harvey J, Arnold Vedlitz, and James DeNardo, "Does Heavy Turnout Help Democrats in Presidential Elections?," *American Political Science Review*, 1986, pp. 1291–1304.

Wittman, Donald, "Candidate Motivation: A Synthesis of Alternative Theories," *American Political Science Review*, 1983, *77*, 142–157.