

The Common Good and Voter Polarization

Chad Kendall and John G. Matsusaka*

May 19, 2022

Abstract

Do voters see democracy entirely in spatial terms, as a trade off of inherently conflicting interests, or do they also view it as a search for the “common good”, as some democracy theorists have long conjectured? We develop an empirical model in which voters have preferences over both common-good and spatial payoffs, and provide a novel method to disentangle the two. Estimating the model on California ballot propositions from 1986 to 2020, we find that 74 percent of voters placed significant weight on the common good, and that partisan polarization roughly doubled among the public over the last decade, mainly due to Democrats drifting to the left.

1 Introduction

Democracy has its roots in two venerable traditions. One, going back at least to Aristotle, sees democracy as a search for the common good, policies that redound to the benefit of all. By involving the people in self-government, the dispersed information of the many helps identify the common good better than decisions made by a

*Department of Finance and Business Economics, Marshall School of Business, University of Southern California, Los Angeles, CA 90089-1422 (e-mail chadkend@marshall.usc.edu and matsusak@usc.edu). Some of the data for this study, originally collected by the Field Research Corporation, were provided by the University of California Data Archive & Technical Assistance (UCDATA). UCDATA is not responsible for the analysis and interpretation of the data appearing here. We thank Matilde Bombardini, Odilon Camara, Nathan Canen, Ernesto Dal Bó, Patrick Francois, Seth Hill, Chris Tausanovitch, Francesco Trebbi, and workshop/conference participants at UC-Berkeley, Université Laval, USC, and the Vancouver School of Economics for valuable input.

small group or single person.¹ A more recent tradition, often called “pluralism,” sees democracy instead as an arena for the resolution of inherently conflicting interests. Competition leads groups to check and balance each other, resulting in collective decisions that reflect compromise and a balancing of conflicting interests.² A tension has always existed between these two traditions, but it seems to have become acute recently with the metastasizing of partisan polarization.³ In a world where the population seems to have divided into irreconcilable camps, one may wonder if the idea of the common good has been lost in a sea of partisanship.

We have little evidence that speaks to this issue, in part because the concepts of common good and spatial interests seldom intersect in empirical work. For example, the foundational evidence on polarization, the NOMINATE scores of Poole and Rosenthal (1985, 1997), assumes purely spatial preferences. Yet in practice, voters are often asked to weigh common-good considerations against heterogeneous (spatial) interests. Consider the following hypothetical ballot measure:

Ballot Measure #1. *To fund levee improvements for flood prevention by assessing an income surtax on the wealthy.*

In this example, voters must weigh a common benefit – preventing the levee from breaking – against their individual cost stemming from the distribution of the tax burden. Existing research has little to say about how voters make this tradeoff – whether they place weight on the common-good component, or are so polarized that they focus only on the spatial consequences.

This paper develops an empirical approach for estimating the weight voters place on the common-good component relative to the spatial component, and applies it to estimate voter preferences from California ballot propositions over the last three decades. With the hypothetical ballot measure above as a stepping off point, we develop a discrete choice framework in which voters choose between two policy alter-

¹The Condorcet Jury Theorem is in this tradition, as is the large theoretical literature on information aggregation (for a review, see Nitzan and Paroush, 2017). Ober (2008) applies theories of information aggregation to explain the political institutions of the most famous ancient democracy, classical Athens; he labels this form of government “epistemic democracy.”

²This approach was embedded in the U. S. Constitution, as highlighted in Federalist Nos. 10 and 51. Other examples include Bentley (1908), Truman (1951), and Becker (1983).

³McCarty (2019) summarizes the literature.

natives, each of which compounds a “common” payoff and an individualized “spatial” payoff. We model the spatial payoff as the distance from an ideal point, following the literature initiated by Poole and Rosenthal (1985, 1997), and refer to a voter’s ideal point as his or her “ideology.”

We model the common good as a payoff that moves all voters’ utilities in the same direction, but for which the optimal direction is uncertain, following the literature on common values associated with the Condorcet Jury Theorem. Voters are assumed to receive informative signals about each policy’s common-good payoff.⁴

In this setup, citizens may vote differently because they have different information about the common-good payoff, because they have different ideologies, or because they put different weights on the two components. Our model nests a pure spatial model at one extreme, and a pure common-good model at the other, thus combining two modeling traditions, one that assumes entirely spatial preferences and the other that assumes common values with incomplete information. Combining common-good and spatial preferences creates a difficult identification problem. The challenge, intuitively, is that if many voters support a particular policy, it could be because it is closer to their ideological positions than the alternative, or because it has a higher common-good payoff. If one had separate data related to both common-good payoffs and spatial positions, the weights could be estimated directly, but such data are seldom available. The paper’s key contribution is to propose a method for separately estimating the common-good and spatial component using only data on voter behavior and characteristics. The basic idea is to identify spatial ideologies by voters’ average positions across issues, and then identify the common-good component by correlations in votes not predicted by spatial preferences. As with all structural models, identification relies on certain untestable assumptions about the economic environment. The key assumption for our model is that centrist policies are more likely than extreme policies, an assumption we microfound with a model of policy setting. We show formally how this assumption is sufficient to disentangle the two preference components and argue why we believe it is plausible.

We estimate the model on 168 California ballot propositions from 1986 to 2020.⁵

⁴We use the term “common good” as a shorthand for a payoff that moves all utilities in the same direction; we do not take a position on whether a policy is normatively “good”.

⁵Previous research that used ballot propositions to estimate preferences include Deacon and

During our period, the propositions spanned a wide range of economic and social issues, including tax increases, tax cuts, primary elections, redistricting, same-sex marriage, capital punishment, and marijuana legalization. An indicator of their importance is the \$3.4 billion spent on proposition campaigns during the period 2000-2020, much more than the \$1.4 billion spent on legislative elections during the same period (Matsusaka, 2020b). Inferring voter preferences from ballot propositions is relatively unexplored and has some advantages over studying candidate elections. In candidate elections, voters choose between two people that make promises on a bundle of issues, promises that they may or may not keep. Consequently, it may be difficult for voters to map candidates into policy outcomes. With ballot propositions, the nature of the choice is unambiguous: voters chose whether to adopt a law that will go into effect exactly as proposed, or to retain the status quo.⁶ Candidate elections are also influenced by valence considerations such as the ability of the contenders, which may obscure the recovery of spatial and common-good preferences over the underlying issues.

Our estimates indicate that voters perceived issues to have substantial common-good as well as spatial components. We find that 74 percent of voters placed a statistically significant weight on the common-good payoff. In terms of magnitude, zeroing out the common-good payoff would have shifted the average voter's probability of supporting a proposal by the same amount as shifting the voter's ideological position 63 percent of the way between the median Democrat and median Republican.

Our ideology estimates indicate that voters were polarized, both in terms of divergence (the overall dispersion of preferences) and party polarization (the tendency of voters to sort ideologically by party), consistent with previous research using other data and methods. Polarization approximately doubled from 1986 to 2020, with most of the increase occurring after 2010. For the post-2012 period, where there is no evidence yet on voter polarization, our estimates imply that polarization grew largely due to movement of Democrats to the left, not movement of Republicans to the right,

Shapiro (1975), an early example; Snyder (1996), using a version of principal components analysis; and Gerber and Lewis (2004), using a purely spatial model. These studies do not investigate polarization, and apart from Deacon and Shapiro (1975) who proxy for the common good using observables, do not incorporate a common-good component.

⁶By law, propositions are required to embrace only a single subject. This may be partly aspirational (Matsusaka and Hasen, 2010), but omnibus proposals are rare.

contrary to what has been found for members of Congress (e.g., McCarty (2019)).

After establishing the baseline results, we explore robustness to key assumptions, and the possibility that our common-good estimates are spuriously capturing other effects. First, we allow signals to be correlated among voters with the same ideological leaning. Second, to allow for the possibility that the common-good component is picking up an unmeasured second (or higher) spatial dimension, we reestimate the model restricting the data to tax and regulation issues, which we expect to be lower-dimensional than the full set of issues. Third, we test for the possibility that a common shock unrelated to information about the common good, such as emotional responses triggered by campaign messaging, causes votes to move in the same direction. We continue to find statistically and economically significant common-good preferences in all three alternative specifications, and that campaign messaging is not driving the results. We also report descriptive evidence relating our common-good estimates to specific issue types to suggest that our estimates align with intuitions about where common-good payoffs are likely to be prevalent.

Our empirical model draws from several streams in the literature. The idea that voters have spatial preferences, and that they can be inferred from variation in votes across individual issues, is in the tradition of the literature following Poole and Rosenthal (1985, 1997). In terms of common-good payoffs, Iaryczower and Shum (2012) develop a model of judges with spatial preferences and private information about a common payoff, which in their context is the legally “correct” decision. Our model contains an important additional layer of complexity – variation in policy positions – that is absent in their setup where it can be reasonably assumed that spatial preferences (pro-defendant or pro-plaintiff) are constant across cases. From Londregan (1990a, 199b), we draw the idea of using *ex ante* theoretical considerations to model the policy proposal process, thereby providing additional information about the policy locations.

Substantively, our findings are related to the literature on “sociotropic” voting (Kinder and Kiewiet, 1979, 1981), the main finding of which is that voters place more weight on overall economic conditions than their own personal conditions. Those findings suggest that voters consider something akin to the common good when evaluating candidates. Our paper also contributes to the literature on political polarization.

Most of that literature has focused on political elites, with the well-known finding that polarization among elites has increased since the 1970s (McCarty et al., 2016). An open question is whether this trend reflects an underlying polarization among voters or is happening independently of voters (Abramowitz and Saunders, 2008; Fiorina and Abrams, 2008). Our reading of the evidence, especially Hill and Tausanovitch (2015), is that voters have not been taking more extreme positions – at least through about 2010 – but appear to have been sorting by party (Gentzkow, 2016; McCarty, 2019). We contribute to this literature by providing arguably cleaner estimates of ideological positions by accounting for the common-good component, and by tapping the largely unexploited pool of information about voter preferences latent in ballot proposition votes. We also extend the evidentiary base on voter ideology into the most recent decade – the most recent existing estimates end in 2012 – introducing the novel finding that polarization has grown significantly since 2010, both in terms of voters taking more extreme positions and party sorting. These results suggest that voters have been following, not leading, their elected representatives in polarizing.

Finally, our paper is related to the literature that attempts to estimate the quality or “valence” of candidates for office. Kendall et al. (2015) and Cruz et al. (2019) estimate voters’ weights on the unobserved valence of candidates in the context of field experiments, using elicited beliefs from survey data to identify the valence component. Buttice and Stone (2012), Beath et al. (2016), and Iaryczower et al. (2020) estimate candidate valence assuming that it can be proxied by observable characteristics such as education, sex, etc. Valence is conceptually similar to a common-good payoff in that it affects the utility of all voters in the same way. However, while these studies can recover common payoffs associated with candidate characteristics, they are unable to tell us if there are common payoffs embedded in the policy choices themselves. One of our contributions is thus to provide estimates about the amount of common-good considerations associated with policy issues. We add to the study of common-good preferences by identifying a key assumption on the issue space that allows the preferences to be inferred without having an observable empirical proxy, thus potentially widening the scope of problems that can be studied.

2 Empirical Model

We develop a two-stage model in which each citizen, $i = 1, 2, \dots, N$, votes on a series of ballot measures or issues, $j = 1, \dots, J$. For each issue, j , a randomly selected voter proposes a new policy, x_j , to replace a randomly selected status quo policy, q_j . Each voter then casts a vote for q_j or x_j . The following subsections present and analyze each stage of the model, beginning with the second stage.

2.1 Voter Decisions

Each voter i has an ideal point (or ideology) θ_i in a one-dimensional policy space. Voters choose between the policy options q_j and x_j in this space. They derive expressive utility from a spatial component and a common-good component. The common-good component is a payoff that moves the utility of all voters in the same direction (given the same information); voters may disagree about the importance of the common good because of heterogeneous information, or because they place different weights on the common-good payoff. Voter i 's utility from voting for $k_j \in \{q_j, x_j\}$ is given by

$$u(k_j) = - \underbrace{\left(k_j - \tilde{\theta}_{ij}\right)^2}_{\text{private}} + w_i \underbrace{V_j(k_j)}_{\text{common}} \quad (1)$$

$V_j(k_j)$ is the (stochastic) common-good component of the policy and the weight, w_i , reflects the extent to which the voter cares about the common good.⁷

We assume that the common-good component exists in a dimension separate from the spatial dimension: $V_j(k_j)$ is a binary random variable drawn from $\{0, 1\}$ with equal probability for both x_j and q_j . In the levee example of the introduction, voters have spatial preferences over the taxes they pay independent of their perceived value of flood protection. Voters learn about the common-good payoff by receiving an informative signal. The difference in the common-good payoff between the proposed and status quo policy is denoted ψ_j , which takes one of three values: $\psi_j \in \Psi \equiv \{-1, 0, 1\}$. Assuming that each possible value of a policy's common-good component is equally likely, the prior on $\psi_j = 0$ is one-half, and the prior on each of the two

⁷In this interpretation, the benefit of the common good is constant across voters, but voters weight the common good heterogeneously. An alternative interpretation is that voters derive heterogeneous benefits from the common good.

extreme differences is one-quarter. For any issue j , common-good considerations matter if $\psi_j \in \{-1, 1\}$, and do not matter if $\psi_j = 0$. Voters receive a signal, $s_{ij} \in S \equiv \{-1, 0, 1\}$ that indicates the true state with probability π_i , and one of the other two states with probability $\frac{1-\pi_i}{2}$. In our main specifications, s_{ij} is distributed independently across voters, but we allow for correlated signals in a robustness check.⁸

For the spatial component, we assume that voters do not know the precise locations of q_j and x_j , but know the midpoint between policies, $m_j \equiv \frac{x_j + q_j}{2}$, the “direction” of the proposed alternative, $\mathcal{D}_j \equiv I(x_t > q_t)$, and the distribution of $x_j - q_j$ (conditional on m_j and \mathcal{D}_j). Empirically, the midpoint and direction are simpler to identify than the individual policy locations because voters to the left of m_j vote for one policy while those to the right of m_j vote for the other, with the direction pinned down by which of these two groups votes for x_j . This assumption is based on the intuition that voters find it easier to learn the direction and midpoint of the choices on the ballot using endorsements by parties and the media than to learn the individual policy locations.

Voter i votes for x_j if it provides higher expected utility than q_j (the tie-breaking rule is inconsequential) conditional on his or her information, $I_{ij} = \{m_j, \mathcal{D}_j, s_{ij}\}$:

$$\begin{aligned} E \left[- \left(x_j - \tilde{\theta}_{ij} \right)^2 + \left(q_j - \tilde{\theta}_{ij} \right)^2 + w_i V(x_j) - w_i V(q_j) | I_{ij} \right] &\geq 0 \\ \iff E \left[(x_j - q_j)(\tilde{\theta}_{ij} - m_j) + w_i (V(x_j) - V(q_j)) | I_{ij} \right] &\geq 0 \\ \iff \alpha_j^d (\tilde{\theta}_{ij} - m_j) + w_i E [V(x_j) - V(q_j) | I_{ij}] &\geq 0 \end{aligned} \quad (2)$$

where $\alpha_j^d \equiv E[(x_j - q_j) | m_j, \mathcal{D}_j = d]$ and $E[V(x_j) - V(q_j) | I_{ij}] = E[\psi_j | s_{ij}]$ is easily calculated using Bayes’ rule (Appendix B provides the calculations).

We assume that voter i ’s ideological position is subject to an idiosyncratic shock for each issue j , $\tilde{\theta}_{ij} = \theta_i + \varepsilon_{ij}$, where ε_{ij} is drawn from a standard normal distribution. The ideological shocks are independent across voters and independent of ψ_j , s_{ij} , q_j , and the proposer’s identity.⁹ Denote the probability that citizen i votes for x_j

⁸Our decisions to use a discrete number of states and signals and to assume fixed priors are driven by identification requirements. See Appendix B for further details.

⁹Having shocks affect ideology produces a similar specification as having shocks affect utility. However, with shocks to ideology, we can obtain approximate estimates of the ideologies in a first

conditional on $\psi_j = \psi$ and direction $\mathcal{D}_j = d$ as $\gamma_{ij}^{\psi d} \equiv Pr(Y_{ij} = 1 | \psi_j = \psi, \mathcal{D}_j = d)$. Then

$$\begin{aligned} \gamma_{ij}^{\psi 1} &= \sum_{s \in S} Pr(s_{ij} = s | \psi) Pr(\alpha_j^1 (\theta_i + \varepsilon_{ij} - m_j) + w_i E[\psi_j | s_{ij} = s] \geq 0) \\ &= \sum_{s \in S} Pr(s_{ij} = s | \psi) \Phi\left(\theta_i - m_j + \frac{w_i}{\alpha_j^1} E[\psi_j | s_{ij} = s]\right); \end{aligned} \quad (3)$$

and

$$\gamma_{ij}^{\psi 0} = \sum_{s \in S} Pr(s_{ij} = s | \psi) \left(1 - \Phi\left(\theta_i - m_j + \frac{w_i}{\alpha_j^0} E[\psi_j | s_{ij} = s]\right)\right). \quad (4)$$

These vote probabilities nest a standard spatial model ($w_i = 0$) as well as a purely common value model akin to Condorcet's Jury Model ($w_i \rightarrow \infty$). Nesting allows the possibility that voters care only about spatial payoffs, so that a common-good component does not exist simply by assumption.

2.2 Policy Setting

Denote the (estimated) distribution of ideologies, $T(\theta)$, and the distribution of possible status quo policies, $Q(q)$. For each issue j , a status quo policy q_j is drawn from Q and, independently, a voter $i = p$ is drawn from $T(\theta)$. The randomly selected proposer chooses the policy alternative x_j by maximizing his or her utility. Because the common-good component of the alternative is exogenous, the proposer sets $x_j = \theta_p$, maximizing his or her spatial payoff.

The policy setting model produces a distribution over x which, together with the assumed distribution for q , determines the distributions over policy midpoints, m , and differences, $x - q$. We denote these distributions, $f^d(m | \theta, \mathcal{D}_j = d)$ and $g^d(x - q | \theta, \mathcal{D}_j = d, m_j)$, respectively, where θ indicates the set of all voter ideologies. The distribution $g^d(x - q | \theta, \mathcal{D}_j = d, m_j)$ determines α_j^d for a given m_j and \mathcal{D}_j , so that it does not need to be separately estimated.

estimation step that is independent of the policy setting model, simplifying the estimation procedure overall. Canen et al. (2020, 2022) take a similar approach.

2.3 Identification: Discussion of Intuition and Key Assumptions

Appendix B provides a formal proof of identification of the model parameters. Here, we provide the basic intuition.

In our model, ideologies are identified as in spatial models – by differences in vote choices across voters on the same issue. The direction of each proposed alternative, $\mathcal{D}_j \equiv I(x_t > q_t)$, is identified by whether voters on one end of the spectrum favor the alternative or status quo. The policy midpoints are identified by differences in vote choices for the same voter across issues, as conventional, and also by the policy-setting part of the model.

Existence of a common-good component is identified by correlation in votes not predicted by a purely spatial model. In a spatial model, a vote that differs from its predicted value is explained as an idiosyncratic shock so that votes are uncorrelated across voters, conditional on the spatial parameters. With a common-good component, votes are correlated through the common-good component and signals. The common-good component influences voters near the midpoint; the common-good weight determines the range of voters around the policy midpoint that are “influenced” by the common-good component. If the weight is zero, then all votes are cast as predicted by the spatial model. As the weight increases, voters around the policy midpoint deviate from their spatial preference and vote according to the common-good payoff instead. While the weight determines the “range” of the common good’s influence, the signal precision determines how correlated votes are within the range of influence. If signals are completely uninformative, then votes are as in the spatial model.

To understand how policy setting is crucial for identification, imagine a landslide election. Theoretically, a landslide could happen for two reasons: (i) the proposal or status quo (and therefore policy midpoint) was ideologically extreme, or (ii) the common-good payoff was extremely large. It might seem that we cannot distinguish the two possibilities without independent information on the location of the policy midpoint or the value of the common-good component. However, note that if extreme midpoints are less likely than centrist midpoints, then it is more likely that voters were swayed by a high common-good payoff than an ideologically extreme policy. A key insight is that the common-good component can be identified by assuming that

centrist midpoints are more likely than extreme midpoints. The policy setting part of our model offers one way to microfound such an assumption, but other processes that generate centrist distributions would work as well. This critical feature of our analysis is ultimately a maintained assumption, but there are intuitive reasons to expect it to hold. In practice, extreme midpoints seem unlikely because (i) an extreme status quo policy would not be allowed to stand for long by the legislature; and (ii) few citizens or organized groups would be willing to bear the substantial cost of proposing a new law if it was so extreme that it had no chance of approval.

While the assumption that policy midpoints are more likely to be centrist than extreme is important, other assumptions about the proposal process are not. In our baseline model, the common-good component of a policy is exogenous; that is, not chosen by the proposer. In Online Appendix C, we develop and estimate a model that relaxes this assumption. In this “unidimensional” model, the proposal chooses a policy parameter that determines both the spatial location and the common-good value. This might represent a situation in which the proposer chooses a highly redistributive tax scheme (left on the the spatial dimension) that implies a high deadweight loss (low common-good payoff). The findings of this model are similar to our baseline model; in particular, it implies that voters place significant weight on the common good.¹⁰

An additional assumption is that the proposer does not account for the possibility that the policy may fail, as would be implied by strategic models of the proposal process (Matsusaka and McCarty, 2001). Estimating a strategic model is feasible, but more complex (e.g., see Canen et al. (2020)) – and, more important, strategic considerations tend to make extreme midpoints even less likely, reinforcing the key feature of our nonstrategic model.

3 Data

Our data pertain to California propositions during 1986-2020. The 168 propositions for which we have data were statutes or constitutional amendments that voters could approve or reject by majority vote. Of these, 116 propositions were new laws proposed

¹⁰We chose to present the version of the model in the main text because it fits the data better. A Vuong test rejects the null of equal fit ($p < 0.001$).

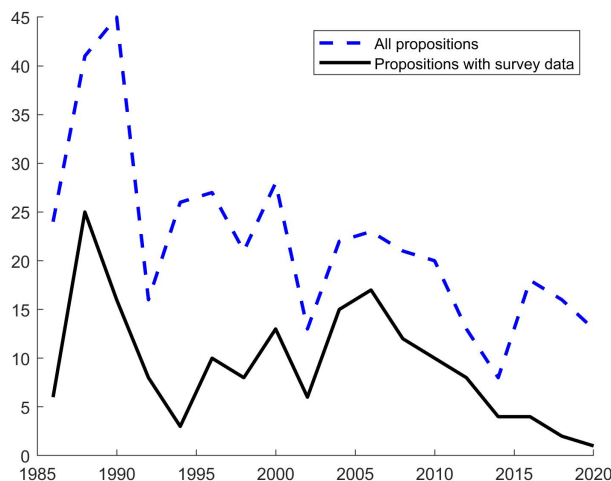
by citizens and three were proposals to repeal existing laws, all of which qualified for the ballot by collecting signatures. With slight abuse of terminology we refer to these together as “initiatives”, even though the latter three are more correctly referred to as “veto referendums.” The remaining 49 propositions were placed on the ballot by the legislature (“legislative proposals”). The range of issues was wide, including tax increases and tax cuts; regulation of insurance companies, farmers, and health providers; social issues such as abortion, same-sex marriage, marijuana legalization, and the death penalty; elections and voting; and government processes, among other things. The propositions spanned the ideological spectrum: some were backed by Democrats, some by Republicans, and some were opposed or supported by both parties.

Figure 1 shows the number of propositions that went to a vote each year, and the number of propositions in our sample (with survey data available). Californians have been voting on issues since the state entered the union in the 19th century; initiatives and veto referendums have been available since 1912. Although California was not the first state to use direct democracy, it has become the leader in using initiatives, and some of its most consequential laws have been the result of citizen initiatives. The number of propositions varies by year, but the flow has been consistent with an average of 22 per two-year electoral cycle. As a result of historical experience, Californians are quite familiar with voting on issues, and can tap a rich array of information sources when deciding how to vote: an official ballot pamphlet containing a nonpartisan analysis from the office of the legislative analyst as well as arguments from proponents and opponents; endorsements and recommendations from politicians, interests groups, and media; and in many cases extensive campaign advertising. As such, we can expect voters to be fairly well informed about most issues, and their votes to reflect their preferences (Lupia, 1994, Lupia and McCubbins, 1998).

For voter preferences on individual propositions we use pre-election survey data from the Field Poll (1986-2012) and Public Policy Institute of California (PPIC) (2010-2020), both well-regarded pollsters in the state.¹¹ The surveys asked voters

¹¹The Field Poll data is available at <https://dlab.berkeley.edu/data-resources/california-polls>. The PPIC data is available at <https://www.ppic.org/data-depot/>. We use the Field Poll for the 2010 and 2012 general elections and PPIC for the 2010 and 2012 primary elections (the Field Poll did not survey the 2010 and 2012 primaries). We did not go back before 1986 because the demographic questions are not readily comparable in the pre- and post-1986 surveys. Data for the 1994 elections,

Figure 1. Number of Propositions by Two-year Electoral Cycle



how they intended to vote on select ballot propositions, their party identification, and demographic information. If there was more than one wave of polling before an election, we use data from the survey nearest to the election, typically a week or two before election day. We use each survey’s recommended sample weights when constructing distributions over policy midpoints, $f^d(m|\boldsymbol{\theta}, \mathcal{D}_j = d)$, and when reporting aggregate results.

We incorporate demographic variables for which comparable data are available across the surveys: categorical dummies for age, education, income, race, and county of residence.¹² For the purposes of party identification, we use a respondent’s self-reported party registration, which in California is simply a designation of which party’s primary the person preferred to participate in (that is, it does not mean the person has any formal connection to the party). After dropping observations with missing data, we have 96,213 responses given by 31,007 respondents. The sur-

the primary elections in 2014, 2016, and 2018, and the general election in 2020 is unavailable. An advantage of survey data is that they do not depend on the turnout decision; turnout shocks, if correlated across voters in particular ways, could potentially bias estimates of the common-good weights.

¹²When response categories varied from survey to survey, we collapsed them into a common set of categories. We do not consider gender because it was not collected for 2002. Counties with fewer than 50 observations were omitted. We exclude the roughly 3 percent of voters that did not identify as White, Black, Asian, or Hispanic, and the 0.4 percent of voters that identified as more than one race

veys cover 168 of the 395 ballot propositions that came before the voters during the period. These propositions were typically higher-profile issues, both in terms of media coverage and campaign spending. For the years in which we have campaign finance data (1998-2016), spending per proposition averaged \$21.5 million for issues that were polled, compared to \$11.7 million for issues that were not polled.

Survey respondents indicated whether they intended to vote for, against, or were undecided about a proposition. Overall, approximately 86 percent indicated that they had an opinion for or against. We omit the 14 percent that were undecided. An alternative approach would be to treat undecided respondents as indifferent between voting for and against, as in Deacon and Shapiro (1975). We chose to drop them because more than half indicated that they had not even heard of the proposition before the survey. Later in the analysis, we make use of this additional data on voters' prior awareness of an issue.

Information on the subject matter of propositions were taken from a database maintained by the Initiative and Referendum Institute (www.iandrinstute.org). Propositions were classified into general categories – taxes, regulation, social issues, elections and voting, government processes, and other – by three coders working independently, as described in Appendix E.

4 Empirical Specification

Because the data include only a handful of votes for each respondent, we construct voter types from respondent characteristics. In particular, we assume $\theta_i = X\beta$ where X consists of the set of observable characteristics described in Section 3, county fixed effects, and time dummies for each four-year period and their interactions with each voter's party registration. We include select interaction terms to limit the number of parameters to be estimated. A consequence of this specification is that the marginal effects of demographics (age, education, etc.) are constant over time.

Absolute ideological movements over time are identified using issues that came to a vote in different years. For example, proposals to require parental notification and a 48-hour waiting period before a minor had an abortion were on the ballot in 2005, 2006, and 2008. There are enough repeat issues to establish links across the

entire period except from 2014 onward, which we therefore treat as a single period (see Appendix F for details).¹³

To identify the common-good weights separately from ideologies requires a form of exclusion restriction in which some observable is a determinant of ideology but not of the weight.¹⁴ We therefore specify $w_i = \exp(W\delta)$, where W includes only demographic observables (excluding party registration, its interactions with time dummies, and county fixed effects). We assume a homogeneous signal precision, but in a robustness check make use of voters' awareness of the issue (section 5.4) to allow the signal precision to differ for voters that were not previously aware of the issue.

For policy setting, we assume that the status quos are drawn from a generalized error distribution, a generalization of the normal distribution, with mean and scale of $\frac{\theta_{max} + \theta_{min}}{2}$ and $\frac{\theta_{max} - \theta_{min}}{2}$, respectively, where $\theta_{max} = \max(\boldsymbol{\theta})$ and $\theta_{min} = \min(\boldsymbol{\theta})$. We set the shape parameter to two, implying a normal distribution, but the findings are similar with other values.

The final parameter vector to be estimated is then $\Theta = \left\{ \{m_j, \mathcal{D}_j\}_{j=1}^J, \beta, \delta, \pi \right\}$. We construct the likelihood of observing a midpoint and a series of votes conditional on this midpoint. Because the distribution of θ can vary from election to election due to both changes in the distribution of likely voters (represented by the sample weights) and changes in preferences over time, the distributions, $f^d(m|\boldsymbol{\theta}, \mathcal{D}_j = d)$ and $g^d(x - q|\boldsymbol{\theta}, \mathcal{D}_j = d, m_j)$, are election specific, $f_e^d(m|\boldsymbol{\theta}, \mathcal{D}_j = d)$ and $g_e^d(x - q|\boldsymbol{\theta}, \mathcal{D}_j = d, m_j)$, $e = 1, \dots, E$. We write $j = 1, \dots, J_e$ for the propositions in election e where $J_1 + J_2 + \dots + J_E = J$.

If one observed the orientation of the policies on each issue j , \mathcal{D}_j , the joint log-likelihood of observing a set of midpoints and their associated votes in the unidimensional model would be given by

¹³With unbounded shocks and sufficient data, the probability that a voter votes Yes or No uniquely determines θ_i , independent of m_j , so that even though the issues change over time, we obtain unbiased estimates of polarization (see also Canen et al. (2022)).

¹⁴Specifically, we require multiple ideologies for the same weight in order to identify the directions of each issue. See the proof in Appendix B.

$$\begin{aligned} \mathcal{L} \left(\left\{ \{ Y_{ij} \}_{i=1}^N \right\}_{j=1}^J ; \Theta \right) = \\ \sum_{e=1}^E \sum_{j=1}^{J_e} \log \left[\sum_{\psi \in \Psi} Pr(\psi) [f_e^{\psi 1}(m_j | \boldsymbol{\theta}, \mathcal{D}_j = 1)^{\mathcal{D}_j} f_e^{\psi 0}(m_j | \boldsymbol{\theta}, \mathcal{D}_j = 0)^{1-\mathcal{D}_j} \right. \\ \left. \prod_{i=1}^N \left(\gamma_{ij}^{\psi 1} \right)^{Y_{ij} \mathcal{D}_j} \left(1 - \gamma_{ij}^{\psi 1} \right)^{(1-Y_{ij}) \mathcal{D}_j} \left(\gamma_{ij}^{\psi 0} \right)^{Y_{ij} (1-\mathcal{D}_j)} \left(1 - \gamma_{ij}^{\psi 0} \right)^{(1-Y_{ij})(1-\mathcal{D}_j)} \right] \end{aligned} \quad (5)$$

The structure of (5) follows from the fact that voters' signals and shocks are assumed to be independent, and that we can write the joint probability of observing each m_j and a set of votes as the product of the marginal probability of observing m_j and the marginal probability of observing the votes conditional on m_j .

We do not know \mathcal{D}_j as assumed in (5), so it must be estimated. Instead of estimating a binary parameter, we calculate the likelihood for each value of \mathcal{D}_j on each issue, j , and then choose the maximum of the two.¹⁵ The likelihood becomes

$$\begin{aligned} \mathcal{L} \left(\left\{ \{ Y_{ij} \}_{i=1}^N \right\}_{j=1}^J ; \Theta \right) = \\ \max_{d \in \{0,1\}} \left\{ \sum_{e=1}^E \sum_{j=1}^{J_e} \log \left[\sum_{\psi \in \Psi} Pr(\psi) [f_e^{\psi 1}(m_j | \boldsymbol{\theta}, \mathcal{D}_j = 1)^{\mathcal{D}_j} f_e^{\psi 0}(m_j | \boldsymbol{\theta}, \mathcal{D}_j = 0)^{1-\mathcal{D}_j} \right. \right. \\ \left. \left. \prod_{i=1}^N \left(\gamma_{ij}^{\psi 1} \right)^{Y_{ij} \mathcal{D}_j} \left(1 - \gamma_{ij}^{\psi 1} \right)^{(1-Y_{ij}) \mathcal{D}_j} \left(\gamma_{ij}^{\psi 0} \right)^{Y_{ij} (1-\mathcal{D}_j)} \left(1 - \gamma_{ij}^{\psi 0} \right)^{(1-Y_{ij})(1-\mathcal{D}_j)} \right] \right\} \end{aligned} \quad (6)$$

We estimate (6) via maximum likelihood using the custom optimization algorithm described in Appendix D. The estimates are consistent for large N and T (e.g. Arellano and Hahn, 2007).

¹⁵See Canen et al. (2022) for another application of this technique in a similar setting.

5 Results

5.1 Voter Estimates

5.1.1 Common-Good Parameters

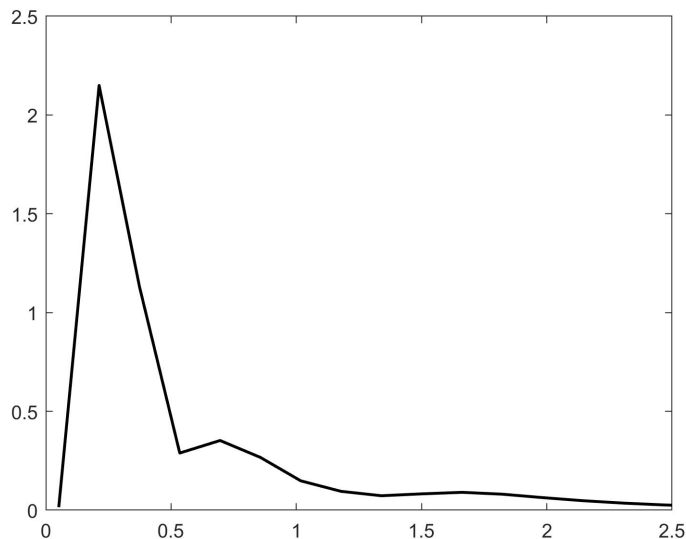
Our first finding is that voters do in fact perceive a common-good component in policy issues. Figure 2 plots the distribution of the estimated common-good weights, w_i , across the entire sample. The mean is 0.68, with standard error 0.17 ($p < 0.001$) and 73 percent of voters have weights that can be differentiated from zero at a significance level of 5 percent. A likelihood ratio test rejects a purely spatial model ($p < 0.001$).

To assess the economic significance of common-good considerations, we calculate the voting probabilities given by (3) and (4), using the estimated signal precision of $\pi = 0.65$ (standard error of 0.02). We then calculate how much the voting probability would change if the common-good component vanished, and how much ideology would have to change to produce an equivalent shift in voting probability. This “equivalent ideological shift” is 0.39 on average, meaning that a common-good voter with a signal would have the same voting probability as a voter 0.39 ideological units away with no weight on the common good. This magnitude seems nontrivial – 63 percent of the distance between the average distance between the median Republican and median Democrat.

Table 1 reports the relations between voter characteristics and (in the left panel) common-good weights and (in the right panel) ideologies. The estimates indicate that common-good weights are lower for older than younger voters, lower for wealthier than poorer voters, and higher for non-white voters (holding constant other attributes); education is not significantly correlated with common good weights. In terms of ideology, where positive values indicate an ideologically conservative position, the estimates imply that older voters are more conservative than younger voters, and highly educated voters are more liberal than less-educated voters, holding constant other attributes; we do not find a significant ideological difference related to income or race/ethnicity.¹⁶

¹⁶One might conjecture that voters with extreme ideologies place less weight on the common good. Figure A2 of Appendix A illustrates some evidence consistent with this conjecture.

Figure 2. Estimated Distribution of Common-Good Weights



Notes: Kernel density estimates of estimated common-good weights.

Table 1. Common Good Weights, Ideology, and Voter Characteristics

	Variable	Estimate		Variable	Estimate
	Age:40-64	-0.50 (0.27)		Age: 40-64	0.12 (0.03)
	Age: 65+	-0.73 (0.36)		Age: 65+	0.20 (0.03)
	College	-0.70 (0.36)		College	-0.09 (0.03)
	College+	-0.75 (0.51)		College+	-0.23 (0.04)
Common good (δ)	Income: 20-60k	-0.74 (0.40)	Ideology (β)	Income: 20-60k	0.03 (0.04)
	Income: >60k	-1.46 (0.45)		Income: >60k	0.07 (0.05)
	Asian	1.93 (0.62)		Asian	-0.00 (0.10)
	Black	1.63 (0.54)		Black	0.03 (0.09)
	Hispanic	1.57 (0.40)		Hispanic	0.05 (0.04)
	Constant	0.41 (0.47)			
	Common good (π)	0.65 (0.02)			

Notes: For β , we do not report the time fixed effects, party coefficients, or their interactions. The omitted categories for β and δ are voters with a high school education or less, voters with incomes below \$20,000/year, voters under the age of 40, whites, and voters not identifying as Republicans or Democrats (β only). Asymptotic standard errors are reported in parentheses. Bold coefficients for δ , β , and π indicate significance at the 5% level or less. For π , we test the one-sided hypothesis that the coefficient is greater than one-third.

5.1.2 Ideological Parameters and Polarization

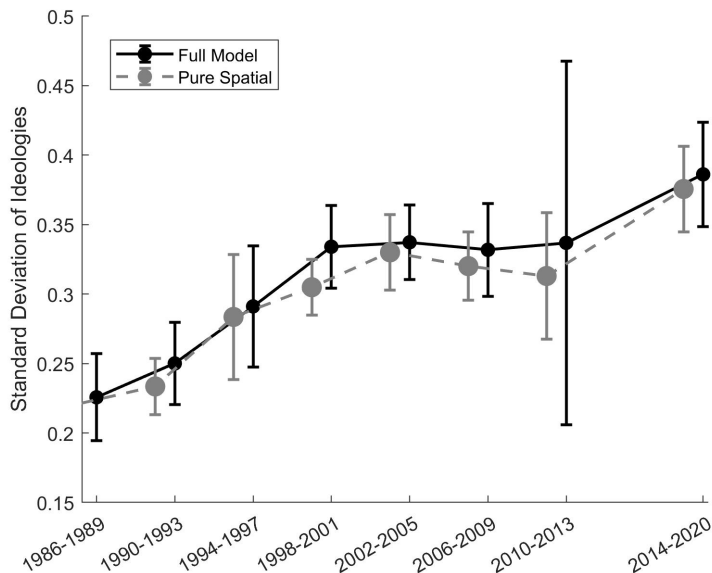
Estimates of polarization among the general public – as opposed to among politicians – are rare (see Hill and Tausanovitch (2015) for estimates and references), and to the best of our knowledge no estimates based on referendum elections exist. Allowing preferences to have a common-good component, we are able to extract spatial preferences that are less at risk of spuriously incorporating common-good effects.

The literature uses two concepts of polarization. The first, which we call “divergence” following Hill and Tausanovitch (2015), is simply the dispersion of ideologies, θ_i . An increase in divergence reflects the replacement of moderate voters by more extreme voters, independent of their party identification. Figure 3 plots the standard deviation of ideologies over time along with estimates from a purely spatial model. The point estimates suggest that the variance of preferences grew from the start of our study period in 1986 until about 2001, remained unchanged for ten years, then diverged sharply in the last decade. Our estimates are somewhat noisy; the standard errors are large enough to preclude confident statements about the details of the time trend, but a tendency toward greater dispersion from the start to the end of the period seems apparent. Comparing across models, it does not seem that omitting the common-good component significantly biases the estimates.

Hill and Tausanovitch (2015) compute voter ideology over the period 1956-2012, using a purely spatial model with data from survey responses to policy questions. Their main conclusion is an absence of a detectable trend in divergence. Examination of their divergence figures suggests there may have been an elevated divergence period from 1956 to 1974, a relatively low period from 1976 to 1996, and a “moderate” period from 1998 to 2012. During the years our samples overlap (1986-2012), they report somewhat noisy evidence of a modest increase in divergence; our findings for this period are fairly similar, giving some reassurance that the two studies are tapping the same things. For the period after their study (2013-2020), we find evidence of a pronounced jump in divergence. If we append our evidence to theirs, the picture for the entire 1956-2020 period can be described as: yearly fluctuations with no evidence of a trend from 1956 until around 1990; a gradual increase in divergence from then on – albeit with annual fluctuations – with a clear increase by 2020.¹⁷

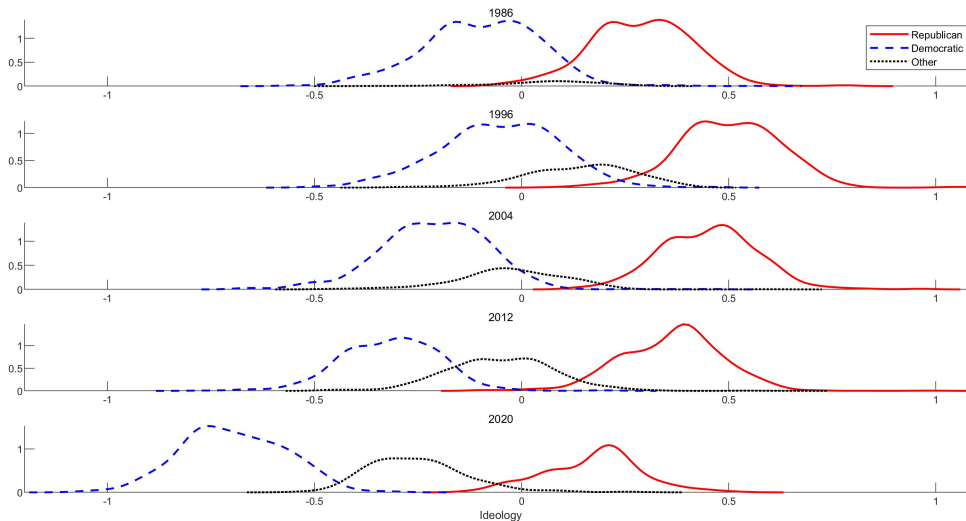
¹⁷We should note that the post-2012 jump appears prior to the 2016 election, so appears to be

Figure 3. Standard Deviation of Ideologies, θ_i , Over Time



Notes: Error bars indicate 95 percent confidence intervals with asymptotic standard errors.

Figure 4. Distribution of Ideologies by Party



Notes: Kernel density estimates of estimated ideologies, broken down by party identification and scaled according to the fraction each type makes up in the population (as obtained from the survey sample weights).

The second measure of polarization, called “sorting,” is the extent to which ideological preferences are correlated with partisan identity. Figure 4 plots the distribution of ideologies at five points in time, distinguishing between voters that identify as Democrats, Republicans, or neither of the two major parties. As one would expect, the figure shows substantial sorting by party in all years. Less obviously, we see that sorting has increased over time and reached an extreme level recently: in 1986, moderate Democrats and Republicans substantially overlapped in ideology, but this overlap had completely vanished by 2020.¹⁸ Furthermore, in contrast to existing evidence on elite polarization in which polarization is mainly attributed to Republicans moving right over time¹⁹, our results suggest that the most recent jump in sorting is due to Democrats moving to the left.

A simple measure of party sorting is the distance between the ideology of the median Democrat and the median Republican, shown in Figure 5. As with divergence, we observe an increase from the start of our period until about 2001, a relatively stable period over the next ten years, and then a large jump in recent years. From 1986 to 2020, polarization by this measure grew by 125 percent. Hill and Tausanovitch (2015, 2018) report a gradual increase in party sorting beginning in the mid-1980s and running through 2012.

The broad conclusion is that polarization among voters appears to have grown during our period of study, both in terms of divergence and party sorting, with a noticeable jump in the last few years. Our findings thus reinforce some existing evidence on polarization, and extend it into recent years. Our evidence casts doubt on the argument that polarization among elected officials has been primarily a response to polarization among the public – to the contrary, the temporal order suggests that polarization among political elites might be fueling polarization among ordinary voters.

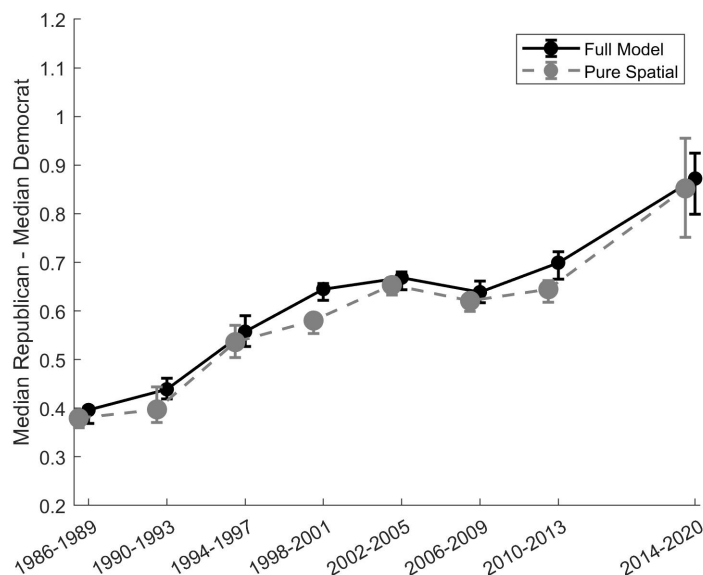
Although polarization by party has attracted the lion’s share of research attention, party identification is not the only potential cleavage point in American society. One popular narrative is that white-collar workers in the cities have gravitated to the

more than a Trump phenomenon.

¹⁸Figure 4 also illustrates an increase in voters that do not identify as Democrats or Republicans – mainly because of a decline in self-identified Republicans.

¹⁹See Canen et al. (2022), for results that dispute this standard view.

Figure 5. Ideological Distance between Median Republican and Median Democrat Over Time



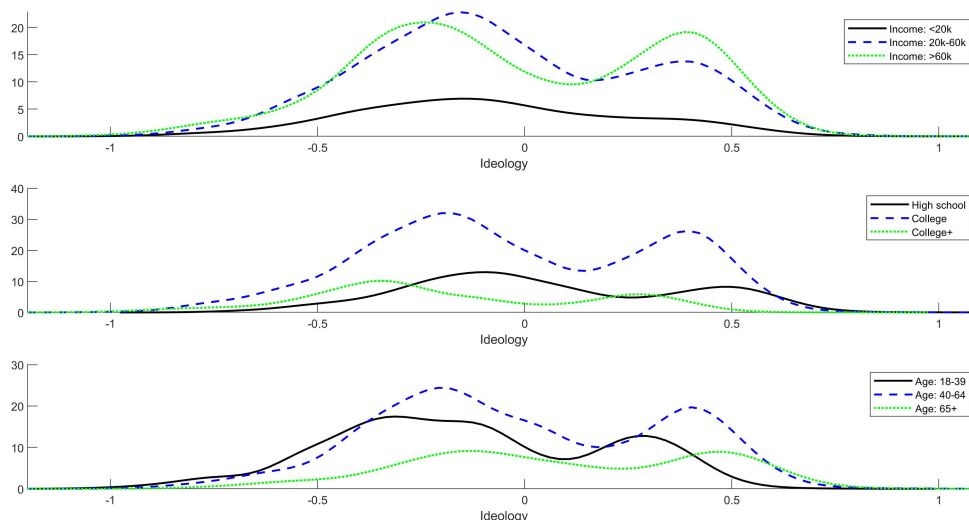
Notes: Errors bars indicate 95 percent confidence intervals with bootstrapped standard errors (200 samples of ideologies drawn from their asymptotic distributions).

Democrats while blue-collar workers in the towns and people living in the countryside have become Republicans. We are interested in the amount of sorting that can be accounted for by demographic factors. We first look at income. There is an ongoing debate about whether the rich or the poor have more influence on policy decisions (Gilens and Page, 2014; Brunner et al., 2013; Lax et al., 2019). The answer matters, of course, only to the extent that the rich and the poor actually have different policy preferences. Figure 6 plots the distribution of ideology by income. Somewhat surprisingly, we find little evidence of polarization by income.

The middle panel of Figure 6 reports the ideological distributions by education. The idea that voters are polarizing along an educational axis has attracted recent public commentary, and is tied to the notion that globalization is creating an environment of high-skilled haves and low-skilled have-nots.²⁰ However, as with income, we find extremists at both ends of the political spectrum and moderates within each educational class.

²⁰For example, <https://www.usatoday.com/story/news/politics/elections/2020/11/24/education-divide-deepens-democrats-worry-future-power/6325025002/>.

Figure 6. Distribution of Ideologies by Income, Education, and Age



Notes: Kernel density estimates of estimated ideologies by income (top panel), education (middle panel), and age (bottom panel). Each is scaled according to the fraction each type makes up in the population (as obtained from the survey sample weights).

The bottom panel of Figure 6 reports the distribution of ideology for three age groups. Again, there is little evidence for polarization by this characteristic. The conjectured tendency of young voters to group on the liberal side of the spectrum does not appear, although this might be due in part to their grouping into a single category (ages 18-39).²¹

5.2 Plausibility of Common-Good Estimates

One of the paper's contributions is identification of a set of plausible assumptions that enable common-good and spatial payoffs to be separately estimated from voter and voting data alone. As seen, the ideology estimates conform to what we expect from previous research, lending some support to the approach. But, because there is no existing literature on common-good payoffs, we cannot compare our common-good estimates with those from other studies in order to assess their plausibility. This

²¹We report the county fixed effects as a choropleth in Appendix A, Figure A1. The expected pattern of more liberal counties along the coast and more conservative counties in the northern interior is clearly present.

section takes an alternative approach, exploring the extent to which the estimated common-good payoffs are correlated with factors that intuition suggests ought to be correlated with them.

5.2.1 Differences Across Individuals

Table 1 reported the correlation of common-good weights with demographic characteristics. While we do not have a strong expectation about most of those correlations, we might expect that it takes time and experience to develop ideology and thus younger voters and less educated voters will have less sharp ideological preferences and place a higher weight on the common-good component.²² Our estimates generally support this idea: common-good weights are highest for the youngest and least-educated voters.

5.2.2 Differences Across Issues

For each proposition, we can calculate the posterior probability of each of the three common-good states, $\psi_j \in \Psi \equiv \{-1, 0, 1\}$, which represent the difference in common-good payoff between the proposal and the status quo. In 85 percent of cases, the estimates place a probability greater than 0.95 on one of the states. Intuitively, if the net common-good payoff was zero, voters viewed the issue as purely spatial. We define the “common-good probability” as the estimated probability that the common-good is relevant for an issue (sum of the posteriors on the negative and positive states). The mean common-good probability is 0.47 with 47 percent of issues having a common-good probability greater than 0.5. Note that the common-good estimates are subjective in the sense that they reflect how voters perceive the issues; they do not necessarily capture the external reality of a common good payoff. It is possible that policies that economists consider to be in the interest of all are not seen that way by the voters. Also, predicting how voters “should” view the payoffs is complicated by the fact that proposals often compound multiple issues, such as imposing an income tax surcharge on the wealthy and using the money to fund children’s hospitals. With these caveats stated, we can nevertheless explore a few possibilities.

²²For example, survey evidence indicates that voters aged 18-29 are the least likely to identify with a party or perceive a difference between the policies of Democratic and Republican candidates (Pew Research Center, 2021).

- *Proposals with high spending by business groups.* We conjecture that propositions with heavy business spending are less likely to have a common good component because they are largely about rent-seeking. Focusing on 92 propositions during 2000-2020 for which we have complete campaign finance data, we find that measures in which business groups spent \$5 million or more had an 11.6 percent lower common-good probability. Focusing just on issues in which business groups spent \$5 million or more in support (i.e., when they were the primary sponsors), the gap was 15.7 percent.²³

- *Bond measures.* Proposals that allocate funds to specific causes might be seen as more distributional in nature than other proposals, but as the levy example suggests, there are cases that go the other way. Bond measures are relatively clean in this respect since they authorize spending on a particular project, without increasing or creating new taxes. Bond measures are used to fund a variety of programs, including highways, school buildings, prisons, parks, and loans to veterans. By focusing on only bond measures, we can compare how voters view different types of spending, independent from the financing mechanism. Overall, bond measures had a 22.3 percent higher common-good probability than other propositions. Water bonds (to build infrastructure for clean water) and education bonds are especially likely to be viewed as having common-good elements. The sample contains only 37 bond measures, but our estimates imply that voters gave water and school bonds a 21.4 percent higher common-good probability than other bonds. The seven water bonds had a common-good probability of 78.7 percent.

- *Taxes.* Taxes are often seen as distributional in nature. We find that proposals concerning taxes had an 8.2 percent lower common-good probability.

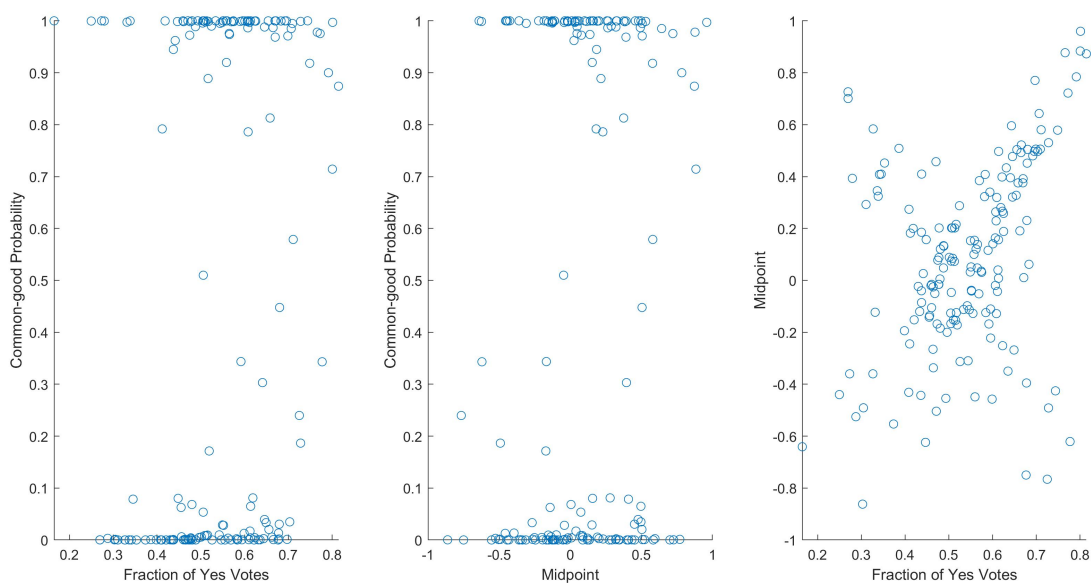
All of these estimates are based on relatively small numbers of observations (essentially the number of propositions), so we offer them only as suggestive evidence.

5.2.3 Common-good Versus Midpoints

Roughly speaking, the model estimates common-good payoffs from comovement in voting. As mentioned above, in principle, voters could all swing in one direction on

²³Campaign contribution numbers were aggregated using data on individual contributions provided by the California Secretary of State.

Figure 7. Common-good Probabilities, Yes Votes, and Midpoints

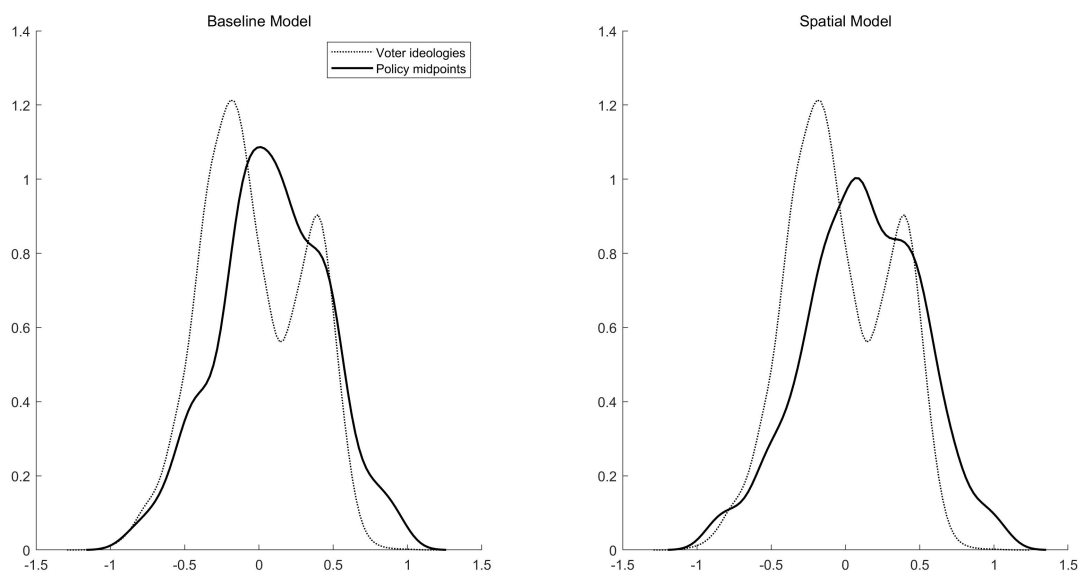


a proposal because it has a large common-good payoff or because it has an extreme midpoint, and we pin down the estimates by assuming that moderate midpoints are more likely to occur than extreme midpoints. Because our identifying assumptions are new to the literature, it is natural to wonder if they are mechanically forcing a high common-good component (and moderate midpoint) on one-sided issues. We offer several pieces of evidence suggesting that this is not the case.

Figure 7 reports scatter plots of the relationships between the common-good probability, Yes votes, and estimated midpoints. The left panel plots the common-good probability against the percentage of Yes votes. We observe high common-good probabilities for close elections, not only for one-sided votes. The middle panel plots the common-good probability against the estimated midpoint, showing that high common-good probabilities occur for a wide range of policy positions. The right panel plots the midpoints against the percentage of votes in favor. We do not see a mechanical relation forcing the midpoint to the middle when the election was one-sided: extreme midpoints occur when the fraction of Yes votes is very large or very small, as in a purely spatial model.

We can also directly compare the estimated midpoints to those from a purely

Figure 8. Distribution of Policy Midpoints for Initiatives and Legislative Propositions



Notes: Kernel density estimates of estimated midpoints for the baseline model, with the distribution of ideologies for reference (left panel). The right panel is for a model which does not include the common-good component or policy setting.

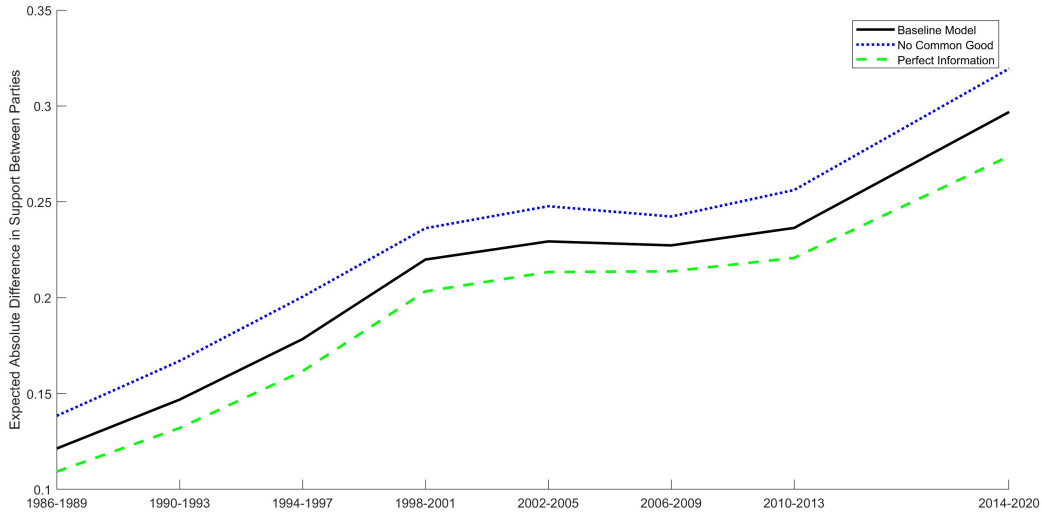
spatial model that does not impose policy setting. Figure 8 plots the estimated midpoints for the full model (left panel) and for a spatial model (right panel), with the density of ideologies for reference. The distributions of midpoints are similar across models. Overall, these results suggest that, while the policy setting information is necessary for identification of the common-good component, it is not mechanically forcing a common-good component on (only) one-sided issues by completely ruling out extreme midpoints.

5.3 Counterfactuals

5.3.1 Partisan Vote Difference

In this section, we report counterfactual exercises that illustrate how common-good considerations affect voting outcomes. Figure A3 in Appendix A, which plots the expected vote shares against the actual vote shares, shows that the model fits the data reasonably well.

Figure 9. Expected Difference in Support between Parties: Counterfactuals



Notes: Plots of the expected absolute difference between parties in fraction of votes in support, averaged over all propositions within each time period under three scenarios: baseline model estimates, a purely spatial model with no weight on the common good, and a model with the estimated common-good weights but perfect voter information.

We first study the expected difference in votes between Republicans and Democrats; this alternative measure of polarization compounds common-good and spatial effects, thereby providing insight into how common-good considerations contribute to, or mitigate, polarization. Figure 9 plots the expected partisan difference in support, averaged over all propositions within each time period (corresponding to the time dummies for ideologies). We show the baseline model for reference, and two counterfactual scenarios: (i) zero weight on the common-good payoff for all voters ($w_i = 0$), and (ii) perfect information about the common good ($\pi = 1$).

The baseline plot shows an increase in polarization over time, consistent with the increase in the difference between party medians observed in Figure 4. Intuitively, common-good concerns should cause people to vote the same way; in the extreme case where voters care only about the common good, they would vote identically on average. Consistent with this intuition, the dotted line when $w_i = 0$ consistently lies above the baseline case. The magnitudes are non-trivial: the partisan difference is 1.8 percentage points higher on average without the common good than in the baseline model, a 8.0 percent increase.

Unlike the common-good weight, which is a preference parameter, the probability of being informed is partially a policy choice. Democracies can influence the flow of information to voters by regulating advertising, media, and campaigning. Debates over campaign regulation often rely on a zero-sum or spatial model of politics, in which publicity for one campaign only hurts the other campaign. In a model with common-good features, however, campaigning potentially increases the quality of public decisions. Without taking a normative stance, we can explore how information about the common-good component feeds back into the partisan nature of voting. The perfect information case allows the maximum impact for common-good considerations. Relative to the case in which citizens place no weight on the common good, partisan differences are 3.3 percentage points lower (a 15% decrease) when citizens have perfect information.

5.3.2 Proposition Passage

Here, we are interested in how common-good considerations affect the type of proposals that voters approve. We start by classifying each proposition as “right-leaning” if a majority of Republicans voted in favor and a majority of Democrats voted against; “left-leaning” if a majority of Democrats voted in favor and a majority of Republicans voted against; and “nonpartisan” otherwise (where “votes” are expected votes under the counterfactual scenario in which voters vote based on ideological considerations alone). Given that Democrats controlled the legislature throughout the sample period, theory suggests citizen initiatives would have come disproportionately from conservatives, while legislative proposals would have been progressive in orientation. We find that 29 percent of initiatives were right-leaning, while only 4 percent of legislative proposals were right-leaning, supporting this theory.

Table 2 provides estimates of the expected passage rates for the two counterfactual scenarios considered in the previous section (no common good and perfect information), along with those from the baseline model and the actual passage rates in the data. The model passage rate predictions are very similar to those in the actual data, again indicating that the model fits the data reasonably well. Right-leaning proposals were much less likely to pass than left-leaning proposals, 36 percent versus 66 percent, consistent with the state’s reputation as being left-leaning (difference significant at

Table 2. Counterfactual Proposition Passage Rates

	All Propositions	Right- leaning	Left- leaning	Nonpartisan
Number of Propositions	168	36	66	66
% Approved, actual	63	39	70	70
% Approved, baseline	64	39	71	71
% Approved, no common good	56	28	58	70
% Approved, perfect information	63	42	68	70

Notes: Propositions were classified as “right-leaning” if a majority of Republicans were in favor and a majority of Democrats were against; “left-leaning” if a majority of Democrats were in favor and a majority of Republicans against; and “nonpartisan” otherwise.

the 1 percent level).

With perfect information the passage rates are relatively unchanged, but with the common-good weight set to zero, the expected passage rates of both left-leaning and right-leaning proposals drop, suggesting that the fact that citizens place weight on the common good results in more policies passing, but does not substantially change the ideological direction of laws.

5.4 Robustness

In this section, we perform several robustness checks, first testing a key assumption of the model and then checking two possible reasons we might spuriously infer a common-good component.

Our model assumes independent signals, but signals may be correlated if voters rely on the same information sources, which is particularly likely among those that share the same party identification. Our first robustness exercise allows for correlated signals within voters of the same partisan leaning (Democrat, Republican, and independent), focusing on the polar opposite case to independent signals - *perfectly* correlated signals.²⁴

Our second and third robustness exercises investigate two possible reasons our

²⁴The likelihood function with correlated signals differs in that it first takes the product over vote probabilities conditional on signals (among voters with the same partisanship) and then sums over signals, instead of the reverse in the independent signals case (6). Proof of identification of the correlated signals model is identical to the case of independent signals as only the structure of the mixtures changes.

model might spuriously infer a common-good component. One possibility is that the ideological space has more than one dimension, and the common-good component is capturing a second spatial dimension. Because different propositions reflect different issues, voters may not perceive them as lying along a single dimension. Although there is no mechanical reason that a second dimension would produce a common-good weight, it seems worth considering the possibility that it is introducing bias in a nonobvious way. If the model is picking up an unmodeled spatial dimension because propositions bundle disparate issues, the common-good weight should substantially decrease or disappear if we focus on only a single type of proposition. The second robustness exercise estimates the model on only tax and regulation issues, which are the most common in the sample (103 propositions); the excluded propositions then pertain mainly to social issues, elections, voting, and government performance.

Another possibility is that votes co-move because citizens are exposed to a common shock through campaign activities. For example, a highly-charged commercial might trigger an emotional response that affects all voters in the same way. This possibility cannot be casually dismissed because experimental evidence from field studies shows that campaigning can change voting decisions (Gerber et al., 20110; Kendall et al., 2015; Rogers and Middleton, 2015), and many California propositions involve heavy campaign spending.²⁵ The baseline model allows for campaign effects through the provision of information; however, campaigning that moves all voters in the same way for affective (non-informational) reasons could induce a spurious common-good effect. To explore this possibility, we rely on the observation that citizens who were unaware of a proposition before being surveyed cannot have been exposed to campaign messaging, so their votes cannot embed a spurious co-movement caused by spending. To identify responses in which voters were unaware of the issue, we use a survey question that asked respondents if they had “seen, read, or heard anything about Proposition X” (or similar language), available for 126 propositions. We then re-estimate the model allowing the signal precision to differ between voter-issue pairs for which voters were aware and unaware. If campaign persuasion is the cause of the estimated common-good effect in the baseline model, then the alternative model should produce an uninformative signal for unaware voter-issues, implying a spatial

²⁵Matusaka (2020b) contains descriptive information and analysis of spending on California propositions.

Table 3. Robustness Estimates

		Baseline	Robustness: Correlated Signals	Robustness: Tax and regulation	Robustness: Vote awareness
Mean w_i		0.68 (0.17)	0.14 (0.03)	0.85 (0.48)	1.19 (0.47)
Mean Equivalent ideological shift		0.39	0.15	0.49	0.38
π	Homogeneous	0.65	1.00	0.67	-
	Aware	-	-	-	0.46
	Unaware	-	-	-	0.59

Notes: Asymptotic standard errors reported for mean w_i in parentheses. Equivalent ideological shift is the change in ideological position that results in the same change in utility as receiving a signal, averaged across the two signals. Each reported value is the mean across voters.

model.

Table 3 reports information related to the estimated common-good component for the baseline model and the three alternatives. A likelihood ratio test rejects, at the 1 percent level, a purely spatial model in favor of a model with the common-good in all three cases.

In the model with perfectly correlated signals, the average weight placed on the common-good decreases, but remains significant at the 1 percent level, indicating that independent signals are not necessary for detection of a common-good component to voter utility functions. A Vuong test rejects the correlated signal model in favor of that with independent signals ($p < 0.001$), suggesting that signals are at least partially independent.

When the model is estimated only for tax and regulation issues, the common-good weights are systematically larger than the baseline model. Due to the smaller sample size, we lose power, but the average common-good weight remains significant at the 5 percent level. Relatedly, if a second ideological dimension is spuriously driving the common-good effect, one might also expect the common-good to be important on issues for which a purely spatial model does not predict well. Using the estimates from a purely spatial model, we find the opposite: there is actually a *positive* correlation between the fit of the spatial model (where higher fit means it predicts more votes) and the expected absolute difference in common-good payoffs across policies, calculated using the estimated posteriors. It does not appear therefore that the common-good component in the baseline model is an artifact of an unmodeled second ideological

dimension.

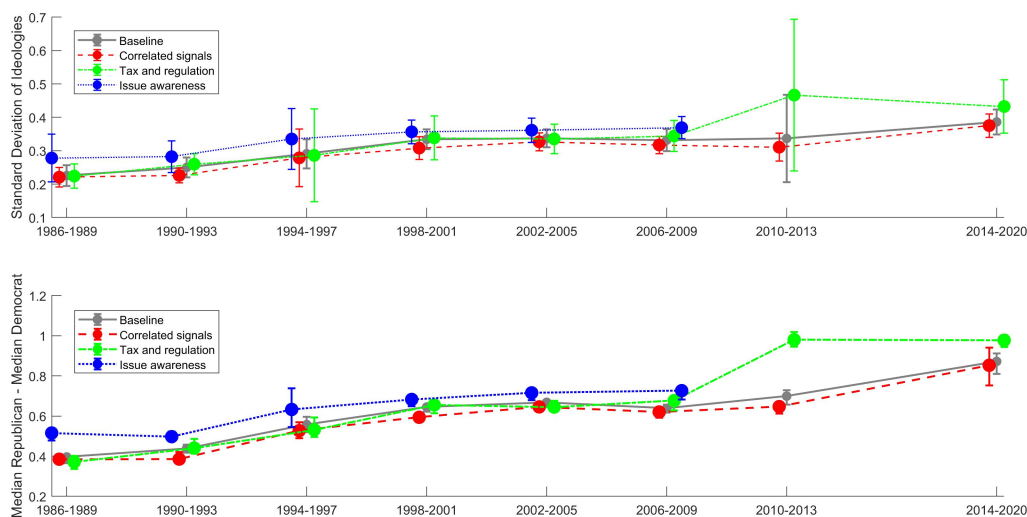
In the third alternative model, in which the signal precision can differ with voter awareness, the average common-good weight is higher than in the baseline model. This is likely an artifact of selection on a subset of issues. The point estimate for the signal precision of unaware voter-issues is statistically significant ($p < 0.001$) and actually larger than that for aware voter-issues, contrary to what would be the case if common-good effects were entirely driven by campaigning. The fact that unaware voters have informative signals suggests that voters may have heterogeneous information even before becoming aware of the issue; for the example discussed in the introduction, voters may have heterogeneous beliefs about the possibility of a flood, and therefore the value of a levy, even before they are exposed to any campaigning.

Figure 10 plots the two polarization measures – divergence in the top panel and partisan sorting in the bottom panel – for the baseline model and its three variants. For the most part, the levels and trends of both measures of polarization are similar to the baseline model for all three alternative specifications. In particular, the levels of polarization are similar when we allow for correlated signals so that ideological polarization does not appear to be driven by differences in partisan beliefs. The one noticeable difference is that when we restrict to tax and regulation issues, both divergence and sorting are higher for the most recent decade (2010-2020). For divergence, the estimates are too imprecise to draw strong conclusions, but the gap for sorting is statistically significant, suggesting that voter opinion has recently pulled apart on taxes and regulation more than other issues. Alternatively, it could be that when ideological positions are estimated across many different types of issues, voters appear more moderate, reflecting inconsistency in their positions across issues.

Lastly, we note that the determinants of the ideologies are stable across specifications (parameters reported in Table A1). With the exception of the correlated signals model, the determinants of the common-good weights are also stable.²⁶

²⁶For completeness, we also report the distributions of ideologies and common-good weights for each of the alternative specifications in Figures A4 through A7. When we restrict to tax and regulation issues, we force voters that don't identify as Republicans or Democrats to have constant ideologies across time because we do not have enough similar issues to bridge ideologies across time periods. In this case, only relative movements in ideologies are identified.

Figure 10. Robustness Polarization Measures



Notes: Divergence and party sorting polarization measures for the baseline model and three alternative specifications: correlated signals, restricted to tax and regulation issues only, and allowing the signal precision to vary with voter awareness. Error bars indicate 95% confidence intervals with asymptotic (top panel) and bootstrapped (bottom panel) standard errors.

6 Discussion

The idea that politics is in part a search for the common good – and not just a zero-sum game between partisans – has a venerable pedigree, running from Aristotle to the U. S. Constitution’s stated goal of providing for “the common defense” and promoting “the general welfare.” It lies at the core of a theoretical literature starting with Condorcet that envisions voting as a way to aggregate information to select policies with the highest common-good payoff. Yet empirical research on voter preferences usually assumes away common-good considerations in favor of a purely spatial model.

Our paper takes a step toward fleshing out the picture by estimating a model in which voters may have both common-good and spatial preferences. Our key insight is that common-good preferences cause votes to co-move, and therefore elections with one-sided voting are more likely to have had common-good features. However, one-sided voting can also be caused by extreme policy midpoints, as purely spatial preferences place most voters on one side of the issue. We show how these possibilities can be separated by a plausible assumption about the distribution of midpoints –

that centrist midpoints are more common than extreme midpoints. We use these ideas to develop a new estimation procedure that allows identification of the weight that voters place on the uncertain common good delivered by a policy. Using data from California ballot measures, we find that common-good payoffs are an economically and statistically significant part of voter preferences.

While our estimates point to a common-good component in voter preferences, they do not elucidate the source of these preferences. Several possibilities seem worth future investigation. One view is that common-good payoffs are technological in nature, hardwired into the issues themselves. For example, the public goods literature studies government projects that provide benefits to all citizens, with national defense a common example. Another view, closely related, is that government policies seek to correct “market failures,” for example, by imposing a Pigouvian tax on an externality such as gasoline consumption. If coupled with a set of compensating transfer payments, such a policy could be Pareto improving, offering benefits to all.

Alternatively, the common-good component could stem from altruism. To the extent that voters are atomistic in a large electorate, they are unlikely to be voting for instrumental reasons. If they vote for expressive reasons, as many voting scholars believe, they may set aside their narrow preferences and take the opportunity to express broader social preferences (Fiorina, 1976; Brennan and Lomasky, 1993). If voters have even a small utility over the well-being of their fellow citizens, the perceived aggregate payoffs to the population at large may end up driving their voting decisions, resulting in a common-good component (McMurray, 2017).

Allowing for a common-good component in voter preferences allows greater confidence in estimates of the spatial component of preferences by removing a potential source of bias. We find that voters were spatially polarized during our sample period, 1986-2020, and polarization grew significantly in the most recent decade. This holds for pure dispersion of preferences (divergence) and for sorting by party (partisan polarization). We find little evidence of significant or growing polarization based on income, education, or age. For the most recent period after 2012, for which there is little polarization evidence, we find a big jump in partisan polarization that is largely the result of Democratic voters shifting to the left, not Republicans shifting to the right.

The existence of a common-good component suggests that it might be possible to reduce polarization by giving voters more information, prompting them to place more weight on common-good considerations. Our counterfactual exercises provide suggestive evidence in favor of this hypothesis, but we are cognizant of the fact that this result rests on the assumption that voters update in a Bayesian manner. More work is needed to determine when information leads to convergence of beliefs (Kendall et al. (2015)) and when it instead leads to an increase in polarization through belief divergence (Baysan, 2022).

References

- [1] Abramowitz, Alan I., and Kyle Saunders. 2008. “Is Polarization a Myth?” *Journal of Politics*, 70 (2): 542-555.
- [2] Allman, Elizabeth S., Catherine Matias, and John A. Rhodes. 2009. “Identifiability of Parameters in Latent Structure Models with Many Observed Variables.” *Annals of Statistics*, 37 (6A): 3099-3132.
- [3] Arellano, Manuel and Jinyong Hahn. 2007. “Understanding Bias in Nonlinear Panel Models: Some Recent Developments.” In *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, eds. Richard Blundell, Whitney Newey, and Torsten Persson, 381-409. Cambridge UK: Cambridge University Press.
- [4] Baysan, Ceren. 2022. “Persistent polarizing effects of information: experimental evidence from Turkey.” *American Economic Review*, forthcoming.
- [5] Beath, Andrew, Fotini Christia, Georgy Egorov, and Ruben Enikolopov. 2016. “Electoral Rules and Political Selection: Evidence from a Field Experiment in Afghanistan.” *Review of Economic Studies*, 83 (3): 932-968.
- [6] Becker, Gary S. 1983. “A Theory of Competition among Pressure Groups for Political Influence.” *Quarterly Journal of Economics*, 98 (3): 371-400.
- [7] Bentley, Arthur F. 1908. *The Process of Government: A Study of Social Pressures*. Chicago, IL: University of Chicago Press.

- [8] Brennan, Geoffrey and Loren Lomasky. 1993. *Democracy and Decision: The Pure Theory of Electoral Preference*. Cambridge, UK: Cambridge University Press.
- [9] Brunner, Eric, Stephen L. Ross, and Ebonya Washington. 2013. “Does Less Income Mean Less Representation?” *American Economic Journal: Economy Policy*, 5 (2): 53-76.
- [10] Buttice, Matthew K., and Walter J. Stone. 2012. “Candidates Matter: Policy and Quality Differences in Congressional Elections.” *Journal of Politics*, 74 (3): 870-887.
- [11] Canen, Nathan, Chad Kendall, and Francesco Trebbi. 2020. “Unbundling Polarization.” *Econometrica*, 88 (3): 1197-1233.
- [12] Canen, Nathan, Chad Kendall, and Francesco Trebbi. 2022. “Political Parties as Drivers of U.S. Polarization: 1927-2018.” Working paper, NBER.
- [13] Cruz, Cesi, Philip Keefer, Julien Labonne, and Francesco Trebbi. 2019. “Making Policies Matter: Voter Responses to Campaign Promises.” Working paper, NBER
- [14] Deacon, Robert T., and Perry Shapiro. 1975. “Private Preference for Collective Goods Revealed Through Voting on Referenda.” *American Economic Review*, 65 (5): 943-955.
- [15] Fiorina, Morris P. 1976. “The Voting Decision: Instrumental and Expressive Aspects.” *Journal of Politics*, 38 (2): 390-415.
- [16] Fiorina, Morris P., and Samuel J. Abrams. 2008. “Political Polarization in the American Public.” *Annual Review of Political Science*, 11: 563-588.
- [17] Gentzkow, Matthew. 2016. “Polarization in 2016.” Working paper, Stanford University.
- [18] Gerber, Alan S., James G. Gimpel, and Donald P. Green. 2011. “How Large and Long-Lasting Are the Persuasive Effects of Televised Campaign Ads? Results from a Randomized Field Experiment.” *American Political Science Review*, 105 (1): 135-150.

- [19] Gerber, Elisabeth R., and Jeffrey B. Lewis. 2004. "Beyond the Median: Voter Preferences, District Heterogeneity, and Political Representation." *Journal of Political Economy*, 112 (6): 1364-1383.
- [20] Gilens, Martin and Benjamin I. Page. 2014. "Testing Theories of American Politics: Elites, Interest Groups, and Average Citizens." *Perspectives on Politics*, 12 (3): 564-581.
- [21] Hill, Seth, and Chris Tausanovitch. 2015. "A Disconnect in Representation? Comparison of Trends in Congressional and Public Polarization." *Journal of Politics*, 77 (4): 1058-1075.
- [22] Hill, Seth, and Chris Tausanovitch. 2018. "Southern Realignment, Party Sorting, and the Polarization of American Primary Electorates." *Public Choice*, 176: 107-132.
- [23] Iaryczower, Matias, and Matthew Shum. 2012. "The Value of Information in the Court: Get It Right, Keep It Tight." *American Economic Review*, 102 (1): 202-237.
- [24] Iaryczower, Matias, Galileu Kim, and Sergio Montero. 2020. "Representation Failure." Working paper, Princeton University and University of Rochester.
- [25] Kendall, Chad, Tommaso Nannicini, and Francesco Trebbi. 2015. "How Do Voters Respond to Information? Evidence from a Randomized Campaign." *American Economic Review*, 105 (1): 322-353.
- [26] Kinder, Donard R. and D. Roderick Kiewiet. 1979. "Economic Discontent and Political Behavior: The Role of Personal Grievances and Collective Economic Judgments in Congressional Voting." *American Journal of Political Science*, 23 (3): 495-527.
- [27] Kinder, Donald R. and D. Roderick Kiewiet. 1981. "Sociotropic Politics: The American Case." *British Journal of Political Science*, 11 (2): 129-161.
- [28] Kingma, Diederick P., and Jimmy Ba. 2015. "Adam: A Method for Stochastic Optimization." *International Conference on Learning Representations*: 1-13.

- [29] Lax, Jeffrey R., Justin H. Phillips, and Adam Zelizer. 2019. "The Party or the Purse? Unequal Representation in the US Senate." *American Political Science Review*, 113 (4): 917-940.
- [30] Londregan, John. 1999a. "Estimating Legislators' Preferred Points." *Political Analysis*, 8 (1): 35-56.
- [31] Londregan, John B. 1999b. *Legislative Institutions and Ideology in Chile*. Cambridge, UK: Cambridge University Press.
- [32] Lupia, Arthur. 1994. "Shortcuts Versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections." *American Political Science Review*, 88 (1): 63-76.
- [33] Lupia, Arthur, and Mathew D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* Cambridge, UK: Cambridge University Press.
- [34] Matsusaka, John G. and Richard L. Hasen. 2010. "Aggressive Enforcement of the Single Subject Rule." *Election Law Journal*, 9 (4): 339-41.
- [35] Matsusaka, John G. 2016. "Ballot Order Effects in Direct Democracy Elections." *Public Choice*, 167: 257-276.
- [36] Matsusaka, John G. 2020a. *Let the People Rule: How Direct Democracy Can Meet the Populist Challenge*. Princeton, NJ: Princeton University Press.
- [37] Matsusaka, John G. 2020. "When Voters Regulate Business: Capture Theory in Reverse?" Working paper, University of Southern California.
- [38] Matsusaka, John G., and Nolan M. McCarty. 2001. "Political Resource Allocation: Benefits and Costs of Voter Initiatives," *Journal of Law, Economics, and Organization*, 17 (2): 413-448.
- [39] McCarty, Nolan M. 2019. *Polarization: What Everyone Needs to Know*. New York, NY: Oxford University Press.

- [40] McCarty, Nolan M., Keith T. Poole, and Howard Rosenthal. 2016. *Polarized America: The Dance of Ideology and Unequal Riches* (second edition). Cambridge, MA: The MIT Press.
- [41] McMurray, Joseph. 2017. "Ideology as Opinion: A Spatial Model of Common-Value Elections." *American Economic Journal: Microeconomics*, 9 (4): 108-140.
- [42] Nitzan, Shmuel, and Jacob Paroush. 2017. "Collective Decision-Making and Jury Theorems." In *The Oxford Handbook of Law and Economics: Vol. 1: Methodology and Concepts*, ed. Francesco Parisi, Chapter 24, Oxford, UK: Oxford University Press.
- [43] Ober, Josiah. 2008. *Democracy and Knowledge: Innovation and Learning in Classical Athens*. Princeton, NJ: Princeton University Press.
- [44] Pew Research Center, "Younger U.S. adults less likely to see big differences between the parties or to feel well represented by them," December 7, 2021, available at: <https://www.pewresearch.org/facttank/2021/12/07/young-u-s-adults-less-likely-to-see-big-differences-between-the-parties-or-to-feel-well-represented-by-them/>.
- [45] Poole, Keith T. and Howard Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science*, 29 (2): 357-384.
- [46] Poole, Keith T. and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting*, New York, NY: Oxford University Press.
- [47] Rogers, Todd and Joel Middleton. 2015. "Are Ballot Initiative Outcomes Influenced by the Campaigns of Independent Groups? A Precinct-Randomized Field Experiment Showing that They Are." *Political Behavior*, 37 (3): 567-592.
- [48] Snyder, James M. Jr. 1996. "Constituency Preferences: California Ballot Propositions, 1974-1990." *Legislative Studies Quarterly*, 21 (4), 463-488.
- [49] Truman, David B. 1951. *The Governmental Process: Political Interests and Public Opinion*. New York: Alfred A. Knopf.

Online Appendices

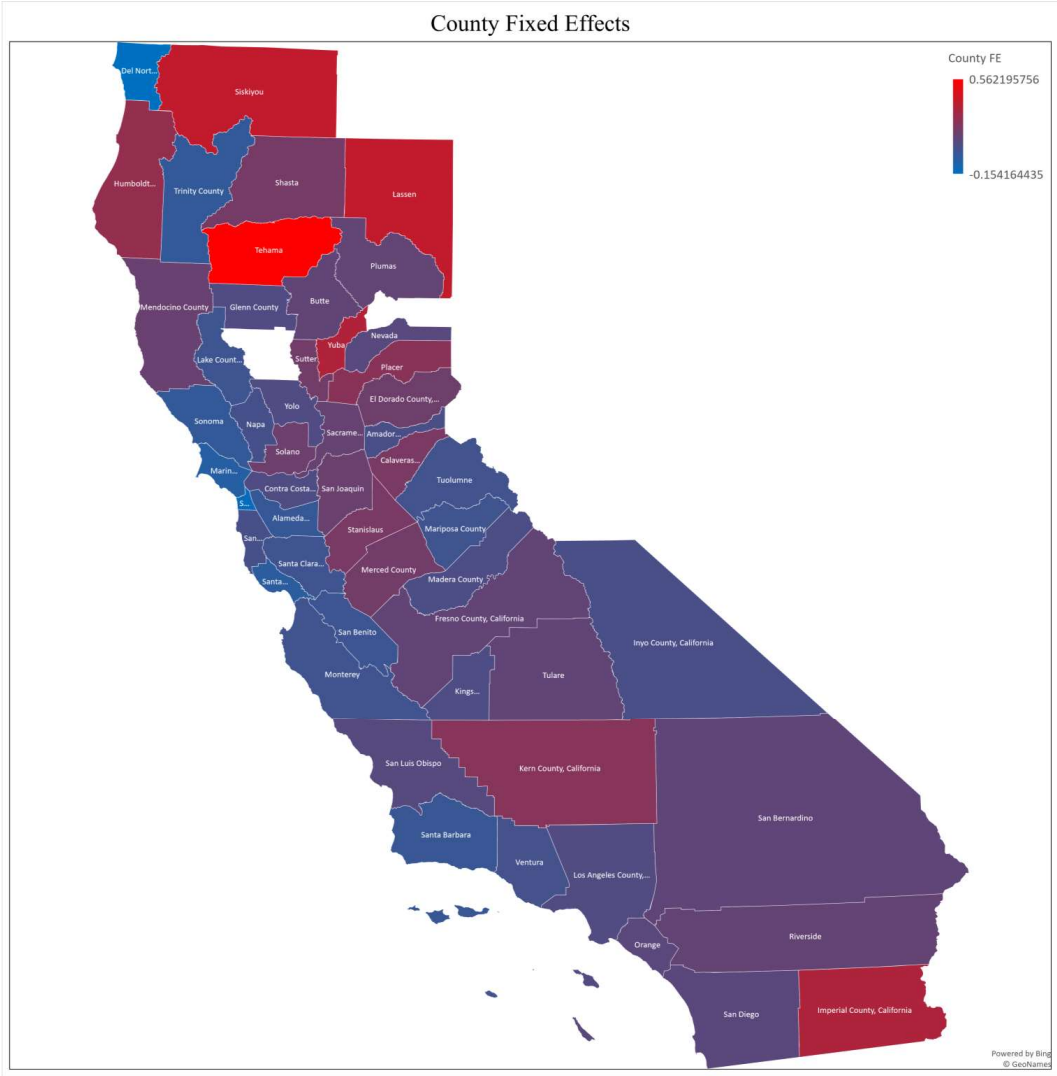
Appendix A: Additional Tables and Figures

Table A1: Robustness: Parameter Estimates

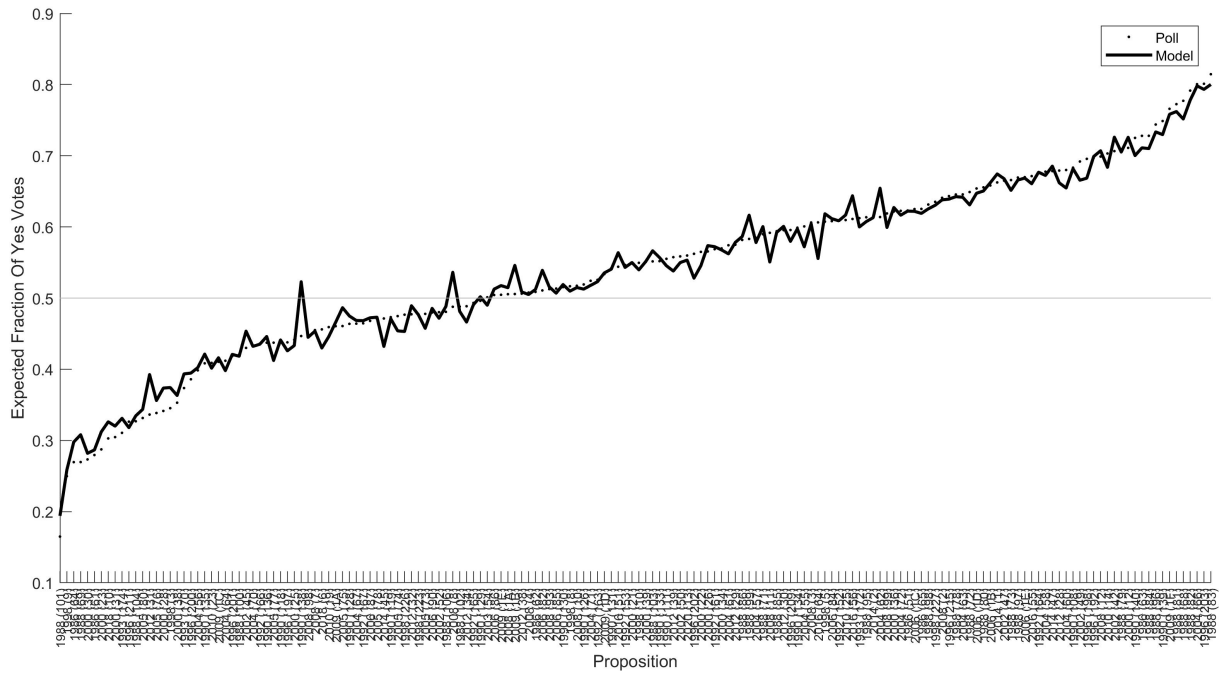
	Correlated Signal	Tax and Regulation	Issue Awareness
Variable	Estimate (s.e.)	Estimate (s.e.)	Estimate (s.e.)
Age: (40-64)	0.34 (0.22)	-1.47 (0.91)	-0.55 (0.29)
Age: (65+)	0.53 (0.20)	-2.12 (0.96)	-1.32 (0.73)
College	0.53 (0.28)	-0.31 (0.63)	-0.60 (0.46)
College+	0.60 (0.24)	-0.49 (0.79)	-1.41 (0.076)
δ Income: 20k-60k	0.63 (0.58)	-0.86 (0.99)	0.08 (0.45)
Income: >60k	0.88 (0.53)	1.66 (1.19)	-0.10 (0.45)
Asian	-1.20 (1.03)	2.74 (1.49)	1.43 (0.95)
Black	-1.51 (1.22)	1.62 (1.29)	1.27 (0.78)
Hispanic	-1.33 (0.76)	1.66 (0.94)	0.81 (0.77)
Constant	-3.26 (0.84)	0.70 (0.65)	1.15 (0.45)
Homogeneous	1.00 (0.03)	0.67 (0.04)	-
π Aware	-	-	0.40 (0.04)
Unaware	-	-	0.56 (0.04)
Age: 40-64	0.12 (0.03)	0.15 (0.06)	0.14 (0.05)
Age: 65+	0.20 (0.03)	0.21 (0.04)	0.23 (0.05)
College	-0.06 (0.04)	-0.05 (0.04)	-0.06 (0.05)
College+	-0.20 (0.04)	-0.23 (0.07)	-0.20 (0.04)
β Income: 20k-60k	0.04 (0.06)	0.08 (0.09)	0.03 (0.06)
Income: >60k	0.08 (0.04)	0.12 (0.07)	0.08 (0.06)
Asian	0.01 (0.04)	-0.07 (0.15)	-0.10 (0.12)
Black	0.02 (0.07)	-0.03 (0.20)	-0.07 (0.11)
Hispanic	0.00 (0.04)	-0.03 (0.08)	-0.02 (0.06)

Notes: Estimated coefficients for three robustness specifications: (1) correlated signals, (2) restricting to tax and regulation issues only, and (3) allowing the signal precision to differ across votes in which the voter was aware and not aware of the issue. For β , we do not report the time fixed effects, party coefficients, or their interactions. The omitted categories for β and δ are voters with a high school education or less, voters with incomes below \$20,000/year, voters under the age of 40, Whites, and voters not identifying as Republicans or Democrats (β only). Asymptotic standard errors are reported in parentheses. Bold coefficients for δ , β , and π indicate significance at the 5% level or less. For π , we test the one-sided hypothesis that the coefficient is greater than one-third.

Figure A1: Estimated Ideologies By County



Notes: Estimates of county-level fixed effect parameters for voter ideologies. The omitted category is the county of Alameda.



Notes: Plot of the expected vote share for each proposition in the data given the model estimates and (estimated) state of the world. The dotted line indicates the actual vote shares in the polling data and the data is sorted in order of these shares.

Figure A2: Correlation Between Ideology and Common-Good Weight

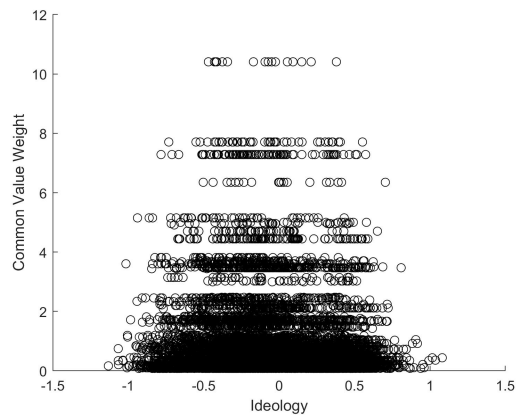
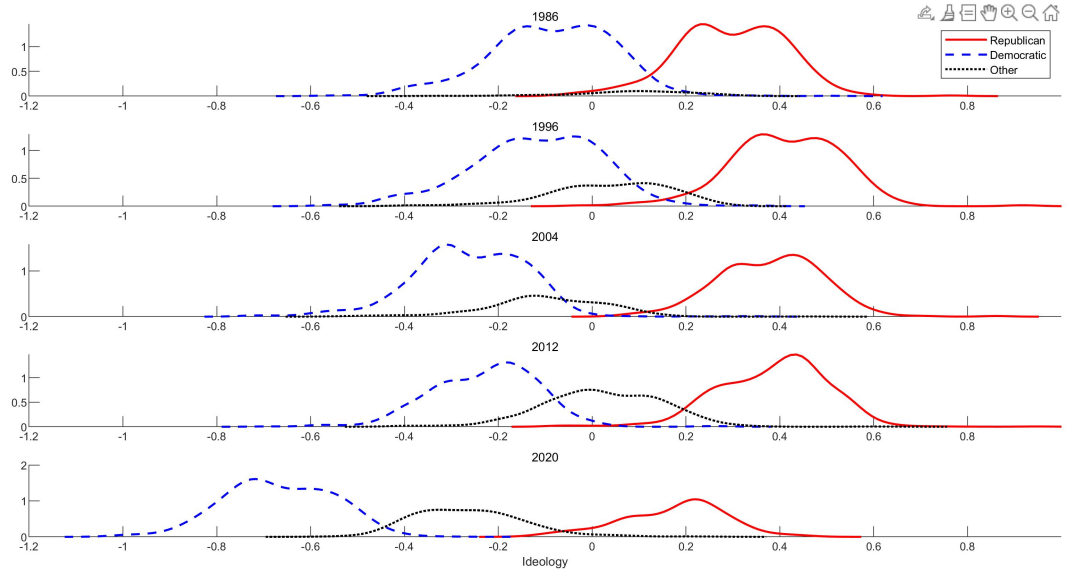
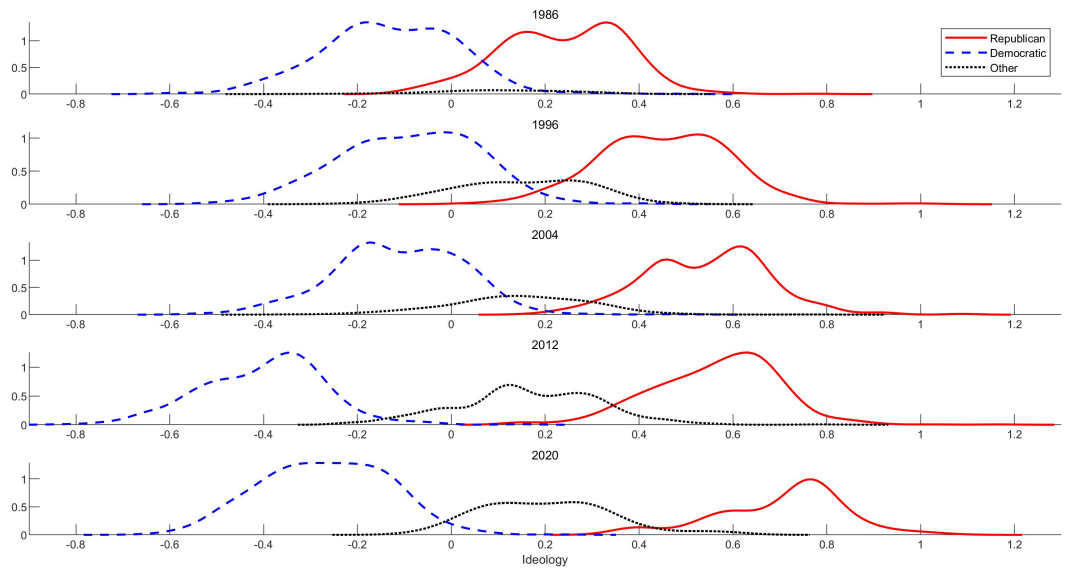


Figure A4: Estimated Distribution of Ideologies (Correlated Signals)



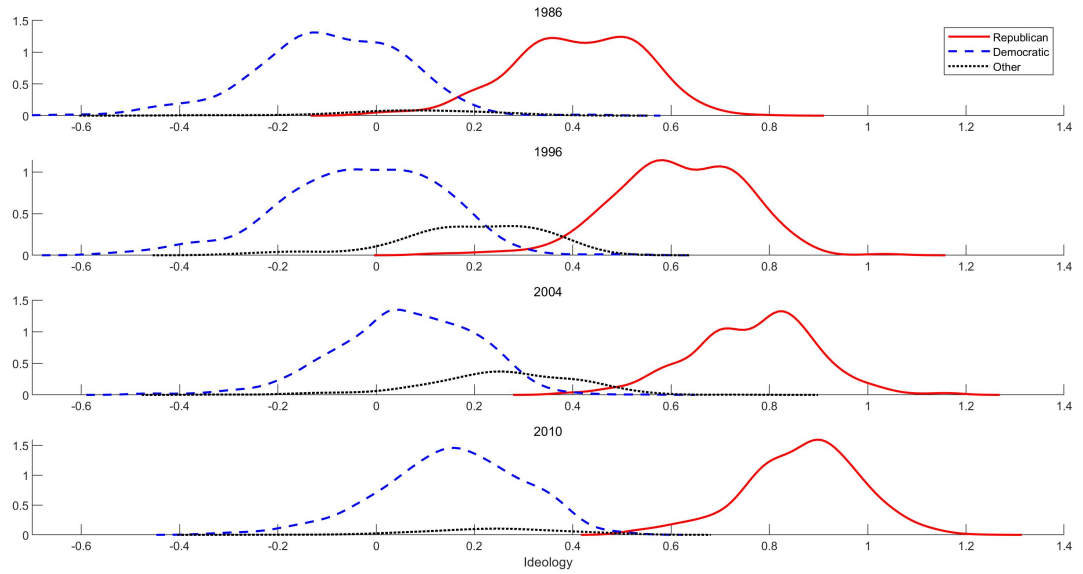
Notes: Kernel density estimates of estimated ideologies broken down by party identification and scaled according to the fraction each type makes up in the population.

Figure A5: Estimated Distribution of Ideologies (Tax and Regulation Issues)



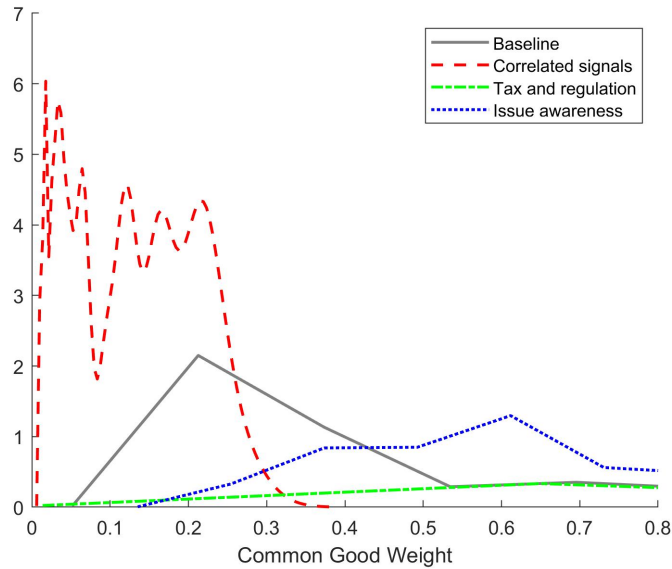
Notes: Kernel density estimates of estimated ideologies broken down by party identification and scaled according to the fraction each type makes up in the population. In this specification, only relative movements in ideologies across party identification are identified because we do not have enough similar issues with which to link ideologies across time.

Figure A6: Estimated Distribution of Ideologies (Issue Awareness)



Notes: Kernel density estimates of estimated ideologies broken down by party identification and scaled according to the fraction each type makes up in the population.

Figure A7: Robustness: Common-Good Weights



Notes: Distributions of the estimated common-good weights under three alternative specifications: restricted to tax and regulation issues only, allowing the signal precision to vary with voter awareness, and assuming a near-uniform distribution for the status quo policies. The baseline model estimates are included for reference.

Appendix B: Proof of Identification

We seek to identify the parameter vector, $\Theta = \left\{ \{w_i, \theta_i\}_{i=1}^N, \{m_j, \mathcal{D}_j\}_{j=1}^J, \pi \right\}$. We provide a constructive proof which shows how identification fails without the information derived from policy setting and then how this information resolves the issue. The proof assumes homogeneous π , but extends immediately to the case in which it is a function of a voter's awareness. We make the following assumptions.

Assumption ID:

1. Voter $i = 1$ has $\theta_1 = 0$.
2. The dataset is rich: there exist at least two voters, i and k , with $\theta_i \neq \theta_k$, $w_i = w_k$, and $s_{ij} = s_{kj}$ for all issues j , almost surely.
3. $\pi \in (\frac{1}{3}, 1]$ and $w_i \in (0, \infty)$ for at least one i

Assumption ID1, together the assumption that ideological shocks have a variance of one, pin down the absolute location and scale of the ideological parameters. These normalizations are standard, necessary assumptions in spatial models.²⁷ Assumption ID2 simplifies identification of the directions of each issue and is satisfied by our empirical dataset.

Assumption ID3 avoids technical complications that arise with parameters at boundaries. With a parameter at a boundary, some of the other parameters may not be identified. For example, if $w_i = 0$ for all i , the model reduces to a purely spatial model so that π cannot be identified.

Throughout the proof, we assume that the conditional distributions of the policy midpoints, m_j , and the difference in policies, $x_j - q_j$ are fixed. In our empirical application they are fixed within each election, but vary across elections due to the changing composition of likely voters. Identification is then obtained within each election.

Preliminaries

We begin by deriving $E[\psi_j | s_{ij}]$ for each signal realization. We do so to show that the expectations, and therefore vote probabilities associated with each signal, maintain a particular ordering for any parameterization of the model. Absent this ordering, the model is not identified because the vote probabilities in the data can then

²⁷Here, one needs to also normalize the scale of the common-good component in order to be able to identify the weight. This normalization appears in the definition of the model.

be associated with multiple possible signals and therefore multiple possible parameter values (see also Iaryczower and Shum (2012)). In particular, if we allow for arbitrary priors, the model is not identifiable because no particular ordering is guaranteed.²⁸

Using Bayes' rule, we have

$$E[\psi_j | s_{ij} = h] = \frac{\frac{1}{4}\pi - \frac{1}{4}\left(\frac{1-\pi}{2}\right)}{\frac{1}{4}\pi + \frac{3}{4}\left(\frac{1-\pi}{2}\right)}$$

$$\begin{aligned} E[\psi_j | s_{ij} = c] &= \frac{\frac{1}{4}\left(\frac{1-\pi}{2}\right) - \frac{1}{4}\left(\frac{1-\pi}{2}\right)}{\frac{1}{2}\pi + \frac{1}{2}\left(\frac{1-\pi}{2}\right)} \\ &= 0 \end{aligned}$$

$$E[\psi_j | s_{ij} = l] = \frac{\frac{1}{4}\left(\frac{1-\pi}{2}\right) - \frac{1}{4}\pi}{\frac{1}{4}\pi + \frac{3}{4}\left(\frac{1-\pi}{2}\right)}$$

Because $\pi > \left(\frac{1-\pi_i}{2}\right)$, $E[\psi_j | s_{ij} = h] > 0$ and $E[\psi_j | s_{ij} = l] < 0$, thus establishing the ordering, $E[\psi_j | s_{ij} = h] > E[\psi_j | s_{ij} = c] > E[\psi_j | s_{ij} = l]$.

Define η_{ij}^{sd} as the probability i votes Yes on issue j conditional on $\mathcal{D}_j = d$ and $s_{ij} = s$. We have

$$\eta_{ij}^{s1} = \Phi\left(\theta_i - m_j + \frac{w_i}{\alpha_j^1} E[\psi_j | s_{ij} = s]\right) \quad (7)$$

and

$$\eta_{ij}^{s0} = 1 - \Phi\left(\theta_i - m_j + \frac{w_i}{\alpha_j^0} E[\psi_j | s_{ij} = s]\right) \quad (8)$$

Given the monotonic relationships between these conditional vote probabilities and the expectations, the vote probabilities are also ordered.

Step 1 (identification of π):

The likelihood in (6) represents a finite mixture (over states) over finite mixtures

²⁸We have also run Monte Carlo simulations with arbitrary priors, confirming this fact.

(over signals). Given fixed priors over states, we can recover all of the parameters of the model from the inner mixtures over signals. The probability that we observe the set of votes across all voters on issue j , \mathbf{Y}_j , conditional on $\mathcal{D}_j = d$ and $\psi_j = \psi$ is given by

$$Pr\left(\{\mathbf{Y}_j|\mathcal{D}_j = d, \psi_j = \psi\}_{i=1}^N\right) = \prod_{i=1}^N \left(\gamma_{ij}^{\psi d}\right)^{Y_{ij}} \left(1 - \gamma_{ij}^{\psi d}\right)^{(1-Y_{ij})} \quad (9)$$

with

$$\gamma_{ij}^{\psi d} = \sum_{s \in \mathcal{S}} Pr(s_{ij} = s | \psi_j = \psi) \eta_{ij}^{sd}$$

(9) represents a standard finite mixture model with mixture probabilities given by $Pr(s_{ij} = s | \psi_j = \psi)$. By standard results, each of the conditional vote probabilities, η_{ij}^{sd} , as well as the mixing parameters, are identified (for example, see Allman et al. (2009)) up to an arbitrary classification of which vote probability in the data is associated with which signal and direction (i.e. the s and d associated with each probability must still be identified).²⁹ The ordering established above, $\eta_{ij}^{hd} > \eta_{ij}^{cd} > \eta_{ij}^{ld}$, ensures the association between the vote probabilities and signals is known independent of the direction. From the mixture probabilities, we immediately obtain π , as the largest of the probabilities. The expected values of the state conditional on each signal are then also known because they depend only on π and exogenous parameters.

Step 2 (identification of θ_i and \mathcal{D}_j):

We first define a normalized vote probability by combining (7) and (8) and then taking the inverse of the monotonic normal cdf:

$$\Phi^{-1}\left(\eta_{ij}^{sd}\right) = (-1)^{1-d} \left(\theta_i - m_j + \frac{w_i}{\alpha_j^d} E[\psi_j | s_{ij} = s]\right) \quad (10)$$

The identification problem is apparent from (10). To see it most readily, assume perfect information so that all voters learn the state perfectly. Because only one state is realized for each issue, one can simultaneously adjust the weights, w_i and the

²⁹It is because of the results on identifiability of finite mixtures that we assume discrete sets of states and signals.

midpoints, m_j , without changing the probabilities. For example, if one increases all of the weights, then for all issues for which the expectation of the state is positive, one can increase the midpoint, while decreasing it for issues for which the expectation of the state is negative.

To use the policy setting information to resolve the identification problem, we begin by using Assumption ID2 to calculate the difference between the vote probabilities of two voters, i and k , who have the same weight, $w_i = w_k$, and signal, s :

$$\Phi^{-1}(\eta_{ij}^{sd}) - \Phi^{-1}(\eta_{kj}^{sd}) = (-1)^{1-d}(\theta_i - \theta_k) \quad (11)$$

The sign of the difference in ideologies, together with the sign of the differences in transformed probabilities, determines the direction. Because we have not yet identified the ideologies, we cannot determine the direction, but we can use this fact to construct the set of issues with the same direction: those for which the sign of the differences in transformed probabilities is the same. Let one set have $\mathcal{D}_j = d$ with $d \in \{0, 1\}$ to be determined, so that the other set has $\mathcal{D}_j = d' = 1 - d$.

We take the expectation of (10) over the proposer identity, conditional on the direction, d :

$$E[\Phi^{-1}(\eta_{ij}^{sd})] = (-1)^{1-d} \left(\theta_i - E[m_j | \mathcal{D}_j = d] + \frac{w_i}{\alpha_j^d} E[\psi_j | s_{ij} = s] \right) \quad (12)$$

Regardless of the direction, the highest probability corresponds to $s_{ij} = h$ and the lowest to $s_{ij} = l$, so that we can solve (12) for w_i using the equation corresponding to the highest probability in the data and substitute it out for that corresponding to the lowest probability. For $d = 1$, we obtain³⁰:

$$\begin{aligned} & E[\Phi^{-1}(\eta_{ij}^{l1})] \\ &= (\theta_i - E[m_j | \mathcal{D}_j = 1]) \left(1 - \frac{E[\psi_j | s_{ij} = l]}{E[\psi_j | s_{ij} = h]} \right) \\ &+ E[\Phi^{-1}(\eta_{ij}^{h1})] \left(\frac{E[\psi_j | s_{ij} = l]}{E[\psi_j | s_{ij} = h]} \right) \end{aligned} \quad (13)$$

³⁰We require $\frac{1}{\alpha_j^d} \neq 0$ for $d \in \{0, 1\}$, $E[\psi_j | s_{ij} = h] \neq 0$, and $E[\psi_j | s_{ij} = l] \neq 0$ which will hold for almost all parameterizations.

Constructing (13) for another voter, k , and taking the difference (noting that because of homogeneous signal qualities, $E[\psi_j|s_{ij} = s] = E[\psi_j|s_{kj} = s]$), we have:

$$\begin{aligned}
& E[\Phi^{-1}(\eta_{ij}^{l1})] - E[\Phi^{-1}(\eta_{kj}^{l1})] \\
= & (\theta_i - \theta_k) \left(1 - \frac{E[\psi_j|s_{ij} = l]}{E[\psi_j|s_{ij} = h]} \right) \\
& + (E[\Phi^{-1}(\eta_{ij}^{h1})] - E[\Phi^{-1}(\eta_{kj}^{h1})]) \left(\frac{E[\psi_j|s_{ij} = l]}{E[\psi_j|s_{ij} = h]} \right)
\end{aligned} \tag{14}$$

If $d = 0$, repeating the above exercise results in

$$\begin{aligned}
& E[\Phi^{-1}(\eta_{ij}^{l0})] - E[\Phi^{-1}(\eta_{kj}^{l0})] \\
= & (\theta_i - \theta_k) \left(\frac{E[\psi_j|s_{ij} = l]}{E[\psi_j|s_{ij} = h]} - 1 \right) \\
& + (E[\Phi^{-1}(\eta_{ij}^{h0})] - E[\Phi^{-1}(\eta_{kj}^{h0})]) \left(\frac{E[\psi_j|s_{ij} = l]}{E[\psi_j|s_{ij} = h]} \right)
\end{aligned} \tag{15}$$

For either direction, either (14) or (15) ensure θ_i is unique for all i because the expectations over normalized vote probabilities come from the data, the expectations of the state are known, and we can set $\theta_i = \theta_1 = 0$ using Assumption ID1. (14) and (15) imply different sets of ideologies, however, θ_i and θ'_i . Note though that, $E[\Phi^{-1}(\eta_{ij}^{s1})] = E[\Phi^{-1}(\eta_{ij}^{s0})]$ for $s \in \{l, h\}$ because the expectations come from the data, so that the equations also imply $\theta_i - \theta_k = -(\theta'_i - \theta'_k)$. From (11), we can then determine the direction, d , and therefore which set of ideologies applies. Assume the left-hand side of (11) is positive. Then, if $d = 1$, we must have $\theta_i > \theta_k$, whereas if $d = 0$, we must have $\theta_i < \theta_k$. A similar argument applies if the left-hand side of (11) is negative. Thus, each θ_i , as well as each \mathcal{D}_j , is identified.

Step 3 (identification of m_j , and w_i):

To identify the weights, take the difference of (12) across the highest and lowest probabilities for voter i :

$$E[\Phi^{-1}(\eta_{ij}^{h1})] - E[\Phi^{-1}(\eta_{ij}^{l1})] = (-1)^{1-d} \frac{w_i}{\alpha_j^d} (E[\psi_j | s_{ij} = h] - E[\psi_j | s_{ij} = l])$$

The expectation, $\frac{1}{\alpha_j^d}$ depends only on the previously identified ideologies and d is known, so that w_i is uniquely determined.

Finally, we can recover m_j from (10). Taking $\mathcal{D}_j = 1$ and $s = h$, and solving for α_j^1 , we have

$$\alpha_j^1 = \frac{w_i}{\Phi^{-1}(\eta_{ij}^{h1}) - \theta_i + m_j} E[\psi_j | s_{ij} = h]$$

Substituting into (10) for another voter, k , we have

$$\Phi^{-1}(\eta_{ik}^{h1}) = \theta_k - m_j + \frac{w_k (\Phi^{-1}(\eta_{ij}^{h1}) - \theta_i + m_j)}{w_i E[\psi_j | s_{ij} = h]} E[\psi_j | s_{kj} = h]$$

which is a linear function of m_j (and known quantities), guaranteeing a unique solution. For $\mathcal{D}_j = 0$, a similar argument applies, so that Θ is identified.³¹

Appendix C: Unidimensional Model

Model

In the unidimensional model, the common-good component exists in the same space as the spatial component. We set $V_j(k_j) = -(k_j - \psi_j)^2$ so that the farther the policy is from some *common* ideal location, ψ_j , the greater the loss. Here the common-good component is innately tied to the spatial location of the policies. For example, an extreme liberal proposition that proposes substantial redistribution may have a deadweight loss associated with the required taxes (and therefore a lower common-good component). We let ψ_j take on three possible values: $\psi_j \in \Psi \equiv \{l, c, h\}$, where $l = \theta_{min}$ corresponds to the spatial position of the voter with the lowest ideology

³¹Technically, we obtain only local identification of the spatial parameters because one can always flip all of the ideologies, directions, policy midpoints, and states through $\theta_i = 0$ without changing the observable probabilities. We deal with this global identification issue in estimation by (if necessary) flipping the estimated results such that Republicans lie right of Democrats.

(leftmost), $c = \theta_{med}$ is that of the voter with a centrist (median) ideology, and $h = \theta_{max}$ is that of the voter with the highest ideology (rightmost).

Voters have a uniform prior over the three states and receive a signal, $s_{ij} \in S \equiv \{l, c, h\}$ that correctly identifies the true state with probability $\pi_i \in [\frac{1}{3}, 1]$, $Pr(s_{ij} = a|\psi_j = a) = \pi_i$ for $a = \{l, c, h\}$. With probability $\frac{1-\pi_i}{2}$, the signal indicates each of the other two states.

As in the baseline model, a voter votes for the alternative, x_j , if

$$\alpha_j^d(\tilde{\theta}_{ij} - m_j) + w_i E[V_j(x_j) - V(q_j)|I_{ij}] \geq 0$$

where, here,

$$E[V_j(x_j) - V(q_j)|I_{ij}] = \alpha_j^d (E[\psi_j|s_{ij}] - m_j)$$

so that the expected distance between policies, α_j^d , drops out of a voter's decision. The conditional vote probabilities become

$$\gamma_{ij}^{\psi^1} = \sum_{s \in S} Pr(s_{ij} = s|\psi) \Phi(\theta_i - m_j + w_i (E[\psi_j|s_{ij} = s] - m_j)) \quad (16)$$

$$\gamma_{ij}^{\psi^0} = \sum_{s \in S} Pr(s_{ij} = s|\psi) (1 - \Phi(\theta_i - m_j + w_i (E[\psi_j|s_{ij} = s] - m_j))) \quad (17)$$

For policy setting, the proposer chooses x_j to maximize his or her utility, leading to $x_j = \frac{\theta_p + w_p E[\psi_j|s_{pj}=s]}{1+w_p}$, where we allow the proposer to be informed through her private signal at the time she sets x_j .³² The likelihood is almost identical to the baseline model except that the distributions of the midpoints now depend on the state of the world (through the proposer's information):

³²Given that the proposer is informed at the time she sets the alternative, x_j , voters receive a noisy signal of ψ_j through their observation of the midpoint between policies, m_j . As updating based on this signal is quite complex; for tractability, we assume voters ignore the signal, basing their votes on their private signals only.

$$\begin{aligned} & \mathcal{L} \left(\left\{ \{ Y_{ij} \}_{i=1}^N \right\}_{j=1}^J ; \Theta \right) = \\ \max_{d \in \{0,1\}} & \left\{ \sum_{e=1}^E \sum_{j=1}^{J_e} \log \left[\sum_{\psi \in \Psi} Pr(\psi) [f_e^{\psi 1}(m_j | \boldsymbol{\theta}, \mathcal{D}_j = 1, \psi_j = \psi)^{\mathcal{D}_j} f_e^{\psi 0}(m_j | \boldsymbol{\theta}, \mathcal{D}_j = 0, \psi_j = \psi)^{1-\mathcal{D}_j} \right. \right. \\ & \left. \left. \prod_{i=1}^N \left(\gamma_{ij}^{\psi 1} \right)^{Y_{ij} \mathcal{D}_j} \left(1 - \gamma_{ij}^{\psi 1} \right)^{(1-Y_{ij}) \mathcal{D}_j} \left(\gamma_{ij}^{\psi 0} \right)^{Y_{ij} (1-\mathcal{D}_j)} \left(1 - \gamma_{ij}^{\psi 0} \right)^{(1-Y_{ij})(1-\mathcal{D}_j)} \right] \right\} \end{aligned}$$

Identification

We begin with the conditional expectations of the state:

$$E[\psi_j | s_{ij} = h] = \frac{\frac{1}{3}\pi\theta_{max} + \frac{1}{3}\left(\frac{1-\pi}{2}\right)(\theta_{med} + \theta_{min})}{\frac{1}{3}\pi + \frac{2}{3}\left(\frac{1-\pi}{2}\right)}$$

$$E[\psi_j | s_{ij} = c] = \frac{\frac{1}{3}\pi\theta_{med} + \frac{1}{3}\left(\frac{1-\pi}{2}\right)(\theta_{max} + \theta_{min})}{\frac{1}{3}\pi + \frac{2}{3}\left(\frac{1-\pi}{2}\right)}$$

$$E[\psi_j | s_{ij} = l] = \frac{\frac{1}{3}\pi\theta_{min} + \frac{1}{3}\left(\frac{1-\pi}{2}\right)(\theta_{max} + \theta_{med})}{\frac{1}{3}\pi + \frac{2}{3}\left(\frac{1-\pi}{2}\right)}$$

These expectations lead to the ordering, $E[\psi_j | s_{ij} = h] > E[\psi_j | s_{ij} = c] > E[\psi_j | s_{ij} = l]$ regardless of the parameters of the model, which is a necessary condition for identification. The vote probabilities, conditional on $\mathcal{D}_j = d$ and $s_{ij} = s$, are:

$$\eta_{ij}^{s1} = \Phi(\theta_i - m_j + w_i(E[\psi_j | s_{ij} = s] - m_j)) \quad (18)$$

and

$$\eta_{ij}^{s0} = 1 - \Phi(\theta_i - m_j + (w_i E[\psi_j | s_{ij} = s] - m_j)) \quad (19)$$

Step 1 (identification of π) is identical to that of the baseline model except that

the signals are identified only once the direction is known ($\eta_{ij}^{hd} > \eta_{ij}^{cd} > \eta_{ij}^{ld}$ if $d = 1$ and $\eta_{ij}^{hd} < \eta_{ij}^{cd} < \eta_{ij}^{ld}$ if $d = 0$). This difference requires the subsequent steps to be different from the baseline case.

Step 2 (identification of w_i):

Combining (18) and (19), and inverting the monotonic function, $\Phi()$, we can define the transformed vote probability,

$$\Phi^{-1}(\eta_{ij}^{sd}) = (-1)^{1-d} (\theta_i - m_j + w_i (E[\psi_j | s_{ij} = s] - m_j)) \quad (20)$$

We next use Assumption ID2 to calculate the difference between the vote probabilities of two voters, i and k , who have the same weight, $w_i = w_k$, and signal, s :

$$\Phi^{-1}(\eta_{ij}^{sd}) - \Phi^{-1}(\eta_{kj}^{sd}) = (-1)^{1-d} (\theta_i - \theta_k) \quad (21)$$

The sign of the difference in ideologies, together with the sign of the differences in transformed probabilities, determines the direction, d . Because we have not yet identified the ideologies, we cannot determine the direction, but we can use this fact to construct the set of issues with the same direction: those for which the sign of the differences in transformed probabilities is the same. Let one set have $\mathcal{D}_j = d$ with $d \in \{0, 1\}$ to be determined, so that the other set has $\mathcal{D}_j = d' = 1 - d$.

Intuitively, we want to compare (20) across signal realizations to separate the ideology from the weight. However, because we don't observe different signal realizations for the same voter on the same issue, we must first construct average probabilities by taking the expectation of (20) over issues (proposer identities and proposer signals). Specifically, we take the expectation over the subset of issues with direction $\mathcal{D}_j = d$ conditional on i receiving signal s and some other voter k receiving signal s' :

$$\begin{aligned} E[\Phi^{-1}(\eta_{ij}^{sd} | s_{kj} = s')] &= (-1)^{1-d} (\theta_i - E[m_j | \mathcal{D}_j = d, s_{ij} = s, s_{kj} = s']) \\ &+ w_i E[\psi_j | s_{ij} = s] \\ &- w_i E[m_j | \mathcal{D}_j = d, s_{ij} = s, s_{kj} = s'] \end{aligned} \quad (22)$$

Note that $E[\psi_j | s_{ij} = s]$ is a constant - i 's expectation of the state given signal s - so

taking its expectation conditional on $s_{kj} = s'$ leaves it unchanged.

The signal corresponding to the highest probability in the data depends on the direction: if $d = 1$, $s_{ij} = h$ corresponds to the highest probability, and otherwise $s_{ij} = l$ does. Thus, we do not know which signal corresponds to the highest probability in the data so let $s_{ij} = H$ denote this unknown signal, and $s_{ij} = L$ denote the signal corresponding to the lowest probability in the data. Condition on voter k receiving the opposite signal (i.e. when using the highest probability in the data for i , average over the cases in which k 's probability is the lowest, and vice versa). Differencing (22) across the two cases, $s_{ij} = H$ and $s_{kj} = L$, and $s_{ij} = L$ and $s_{kj} = H$, we have

$$\begin{aligned}
& E [\Phi^{-1} (\eta_{ij}^{Hd} | s_{kj} = L)] - E [\Phi^{-1} (\eta_{ij}^{Ld} | s_{kj} = H)] \\
&= (-1)^{1-d} (1 + w_i) (E [m_j | \mathcal{D}_j = d, s_{ij} = L, s_{kj} = H] - E [m_j | \mathcal{D}_j = d, s_{ij} = H, s_{kj} = L]) \\
&+ (-1)^{1-d} w_i (E [\psi_j | s_{ij} = H] - E [\psi_j | s_{ij} = L]) \\
&= (-1)^{1-d} w_i (E [\psi_j | s_{ij} = H] - E [\psi_j | s_{ij} = L]) \tag{23}
\end{aligned}$$

where the last equality comes from the fact that the two expectations over the midpoint are the same, because it does not matter which voter receives which signal. Independent of d , the right-hand side of (23) is $w_i (E [\psi_j | s_{ij} = h] - E [\psi_j | s_{ij} = l])$, so that, because the expectations of the state are known and the left-hand side comes from the data, each w_i is uniquely determined.

Step 3 (identification of \mathcal{D}_j , m_j , and θ_i):

We now recover the two possible midpoints (due to the unknown direction) on each issue from (20). Denote the two possible values, m_j^d , $d \in \{0, 1\}$. For voter i , we have

$$\Phi^{-1} (\eta_{ij}^{s1}) = \theta_i - m_j^1 + w_i (E [\psi_j | s_{ij} = s] - m_j^1)$$

and

$$\Phi^{-1} (\eta_{ij}^{s0}) = -\theta_i + m_j^0 - w_i (E [\psi_j | s_{ij} = s] - m_j^0)$$

Noting that $\Phi^{-1} (\eta_{ij}^{s1}) = \Phi^{-1} (\eta_{ij}^{s0})$ comes from the data, we can combine the two

equations to write

$$m_j^1 - m_j^0 = \frac{2\theta_i + 2w_i E[\psi_j | s_{ij} = s]}{1 + w_i} \quad (24)$$

(24) implies a unique value of each m_j because it cannot hold for all voters i on a given issue j given independent signal realizations across issues: with $\theta_1 = 0$ from Assumption ID1, we must have $\frac{2w_1 E[\psi_j | s_{0j} = s]}{1 + w_1} = \frac{2\theta_i + 2w_i E[\psi_j | s_{ij} = s']}{1 + w_i}$ on each issue, which, for fixed θ_i , cannot hold for all s, s' pairs. Given m_j unique, $\mathcal{D}_j = d$ is also unique because (20) must hold for $\theta_1 = 0$. θ_i is then unique from (20) for each voter i . Finally, given unique ideologies for two voters, θ_i and θ_k , $\mathcal{D}_j = d$ for all issues j is determined by (21), pinning down each θ_i and m_j , and thus completing the identification of Θ .

Results

The average common-good weight is 1.34, higher than in the baseline model and statistically significant at the 5 percent level. The distribution over weights is provided in Figure C1. 94 percent of voters have a common-good weight that is significantly different from zero at the 5 percent level. These results indicate that particular assumptions are not driving our finding of a common-good component to voter's utility functions.³³

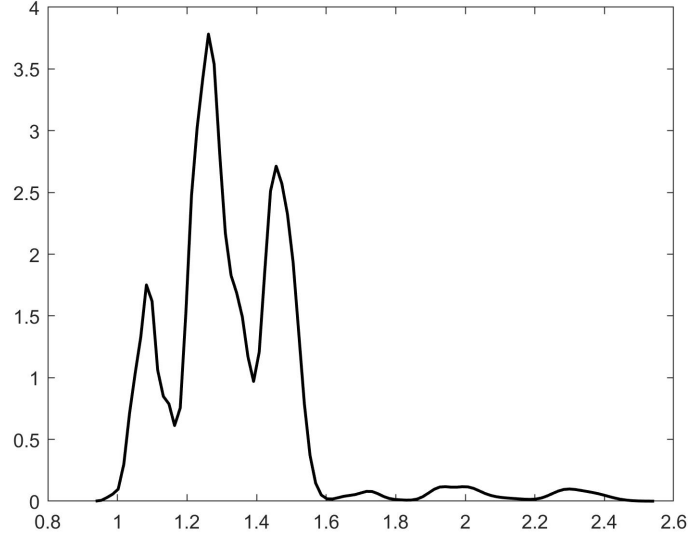
We report the parameter estimates in Table C1 below. We omit the graphs of ideologies and polarization but note that the results are extremely similar to the baseline case.

Appendix D: Estimation Details

The likelihood given in (6) is highly non-convex and therefore poses difficulty for standard estimation procedures, including the expectation-maximization (EM) algorithm that is commonly used for estimating finite mixture models. In particular, we found

³³We observe a significantly positive correlation (0.14; $p = 0.04$) in the common-good components of issues across models. The correlation is somewhat reduced by the fact that the baseline model allows an issue to have zero common-good component whereas the unidimensional model can do so only in non-generic cases in which the midpoint and state are identical. In 75% of cases in which the baseline model predicts a non-zero common-good component, the unidimensional model predicts the same sign for it.

Figure C1: Estimated Distributions of Common-Good Weights (Unidimensional)



Notes: Kernel density estimates of estimated common-good weights.

Table C1: Parameter Estimates for the Unidimensional Model

	Variable	Estimate		Variable	Estimate
	Age:40-64	0.04 (0.13)		Age: 40-64	0.12 (0.03)
	Age: 65+	0.01 (0.15)		Age: 65+	0.20 (0.03)
	College	0.10 (0.16)		College	-0.05 (0.04)
	College+	-0.05 (0.18)		College+	-0.20 (0.04)
Common good (δ)	Income: 20-60k	-0.02 (0.23)	Ideology (β)	Income: 20-60k	0.04 (0.05)
	Income: >60k	-0.19 (0.27)		Income: >60k	0.06 (0.05)
	Asian	0.46 (0.29)		Asian	0.06 (0.05)
	Black	0.02 (0.36)		Black	0.02 (0.07)
	Hispanic	-0.06 (0.16)		Hispanic	-0.00 (0.04)
	Constant	0.28 (0.48)			
	Common good (π)	0.41 (0.12)			

Notes: For β , we do not report the time fixed effects, party coefficients, or their interactions. The omitted categories for β and δ are voters with a high school education or less, voters with incomes below \$20,000/year, voters under the age of 40, whites, and voters not identifying as Republicans or Democrats (β only). Asymptotic standard errors are reported in parentheses. Bold coefficients for δ , β , and π indicate significance at the 5% level or less. For π , we test the one-sided hypothesis that the coefficient is greater than one-third.

that the likelihood function is highly non-monotonic in the policy midpoint (holding the other parameters fixed). After extensive experimentation, we discovered that a version of steepest descent that incorporates momentum in the gradient (Adam et al. (2015)) proved much more efficient and robust. Its one drawback is that it is only defined for unconstrained optimization problems, but this drawback is more than compensated for in terms of speed and robustness.³⁴ Given non-monotonicity in the policy midpoints, we developed the following multi-stage estimation procedure:

1. Begin with a spatial model without the common good. Estimate the ideological parameters, β and midpoints, m_j , as follows:
 - (a) Estimate the model without policy setting. We begin with a global search over 108 randomized starting points for which we calculate the likelihood and take the best 72. For each of the best 72, we perform a grid search over the policy midpoints for each issue. Then, for the resulting 12 best parameters sets, we iterate between using Adam and a policy midpoint grid search until convergence. Convergence is achieved when (i) the midpoint grid search did not change any of the T policy midpoints, and (ii) the Adam algorithm converged (change in the likelihood is less than 0.01 and the infinity norm of the gradient is less than one).
 - (b) For the best parameter set resulting from step a), we re-estimate the model imposing policy setting. Within each estimation loop, we calculate the distributions for the policy midpoints, $f^d(m|\boldsymbol{\theta}, \mathcal{D}_j = d)$, and expected policy distance parameters, α_j^d , for each election for the current ideological parameters. In this estimation step, we run Adam until either convergence is achieved or 500 iterations have completed.
2. Beginning from the ideological parameters estimated with the spatial model, estimate the full model. As for the spatial model, we do this in two steps:
 - (a) Holding the ideological parameters and the policy setting information (i.e. the distributions and expected distances they imply) fixed, re-estimate the policy midpoints, as well as the common-good parameters. The procedure

³⁴In particular, running Matlab’s standard unconstrained or constrained optimizers, `fminunc` and `fmincon`, proved futile. The algorithms rarely converged to the same maximum, and the parameters varied significantly across runs. Moreover, the maxima we found with the estimation procedure outlined here were significantly larger.

is as in step 1a) above.

- (b) As in step 2b) above, re-estimate the model allowing all of the parameters to change.³⁵

With this estimation procedure, we are able to robustly estimate the likelihood function, obtaining very similar parameter estimates over several runs with different (random) parameter initializations.

Appendix E: Issue Classification

The subject matter of each ballot proposition was classified by three researchers: a coauthor of this paper, a finance PhD student with a law degree, and a public policy PhD student. Each classifier was given a list of ballot propositions together with a short description of each proposal, drawn from a database maintained by the Initiative and Referendum Institute, and a classification rubric (below). The rubric contains five broad categories and a residual “other” category. Each researcher classified each proposition as relating to one or more issues. The classifiers were in complete agreement on 74 percent of the measures, and there was a majority view on 97 percent. For 3 percent, there was no consensus.

Rubric

Each proposition is assigned to one or more of the following categories:

- (E) Elections, voting, campaigns, redistricting, term limits, recall, initiative and referendum
- (G) Government processes: procedures for budget approval, civil service reform, organization of legislature, operation of administrative agencies, legislator pay, operation of courts
- (O) Other: Issues not elsewhere classified
- (R) Regulation of business and labor markets
- (S) Social issues: abortion, civil rights, crime and punishment, gay rights, marriage, race, animal rights, drug legalization

³⁵Because we must recalculate the policy setting information within this estimation loop, each iteration is quite time-consuming so we stop after 500 iterations even if convergence is not achieved. The parameter values change very little in this final step so that even if convergence is not achieved, they are close to optimal.

- (T) Taxes, government spending, government borrowing (including education)
- Examples from most recent election (November 2020), with proposed classifications:

Table D1: Example Issue Classifications for the November 2020 General Election

Prop	Description	Classification
14	\$5.5 billion bond issue for stem cell research	T
15	Removes limits on property tax assessment increases for property owned by businesses	T
16	Removes prohibition on government using racial preferences in college admissions and hiring	S
17	Restores voting rights to felons	E
18	Allows 17-year-olds to vote	E
19	Allows disabled elderly homeowners to transfer their property tax exemption to a new home	T
20	Restricts parole for certain offenses	S
21	Allows local governments to control rents, overriding state controls	R
22	Allows rideshare workers to be employed as independent contractors	R
23	Requires physician during kidney dialysis treatment at corporate facilities	R
24	Allows consumers to restrict sale of their digital information	R
25	Eliminates bail payments	S

Our classification scheme resulted in 83 propositions classified as tax issues, 29 classified as regulation issues, 11 classified as government issues, 20 classified as social issues, 25 classified as election issues, and 3 classified as other issues. 6 propositions were not classified and 10 were classified into two categories.

Appendix F: Repeated Issues

In order to compare the ideology of a voter across time, we must observe his or her vote on the same issue in different years. Fortunately, similar ballot propositions do occur repeatedly, allowing us to link together time periods. Table F1 provides a complete list of the pairs or triplets of issues that together link ideologies across time.

Table E1: Repeated Ballot Propositions

Year	Proposition	Description
2005	73	Abortion: parental notification and 4-8 hour waiting period for minor to have abortion
2006	85	
2008	4	
1998	226	Union dues: prohibits use of union dues for political purposes without member consent
2005	75	
2012	32	
2010	19	Legalizes marijuana
2016	64	
2000	22	Defines marriages as solely between one man and one woman
2008	8	
1988	106	Limits attorney contingency fees to 15% (prop 202) and 15% to 25% (prop 106)
1996	202	
1996	198	Allows for open primaries
2004	62	
2010	14	
1990	119	Citizen redistricting commission: 12 members selected by retired judges (prop 119) or 14 members selected randomly (prop 20)
2008	11	
1988	99	Tobacco tax: increase (prop 99) and subsequent repeal (prop 28)
2000	28	