Measuring District-Level Economic and Moral Ideology

Christopher Warshaw
cwarshaw@stanford.edu

Jonathan Rodden
jrodden@stanford.edu

Department of Political Science
Stanford University

December 13, 2009

Abstract: Due to insufficient sample sizes in national surveys, strikingly little is known about issue preferences at the level of Congressional districts in the United States, and existing studies focus only on a single issue dimension. This paper takes a first step toward filling this gap by developing a hierarchical Bayesian model to combine survey and census data to estimate the median of each congressional district's economic and social preferences. Our estimates generate an intuitive new two-dimensional map of American political geography with a variety of substantive applications in American Politics. Moreover, our methodological approach could help scholars in both American and Comparative politics combine surveys and census data to develop estimates for the preferences of sub-national geographic units to test formal models of political economy.

Introduction

The aggregation of voters' preferences into policy lies at the heart of democracy. In countries like the United States, where legislative representation is based on plurality elections in single-member constituencies, it is crucial to understand how those preferences are distributed across voting districts in order to understand whether elected officials, and ultimately the policies they enact, reflect the will of voters. Answers to some basic questions in the study of American politics have remained elusive because of the difficulty of obtaining reliable estimates of district-level policy preferences. For instance, why does the picture of American politics drawn by studies of Congressional roll-call voting look so different from the picture drawn by public opinion research? Does the asymmetric distribution of Democrats and Republicans across Congressional districts lead to bias in the transformation of preferences to policies? Why do some districts regularly split their votes between Congressional and presidential elections?

In spite of the substantive importance of good measures of voter preferences at various geographic scales, we know strikingly little about the ideology of American voters at the district level. Previous empirical work has been stymied by the fact that the sample size in national surveys is generally too small to make inferences at the district-level, and empirical research has adopted the strategy of employing the district-level presidential vote as a catch-all proxy for district ideology.

Perhaps the most crucial weakness of this strategy is its failure to account for the fact that elections and legislative battles cannot always be characterized by a single issue dimension. For example, the 2009 debate over health care reform appeared on first glance to pit economic liberals against economic conservatives. Yet the interjection of the very salient issue of abortion coverage threatened to dissolve the Democrats' tenuous coalition, and a coalition of socially

conservative Democrats forced the reluctant Democratic leadership to allow a vote on the so-called Stupak amendment, which obtained support from all Republicans and a substantial number of socially conservative Democrats. This episode demonstrated that while roll-call votes on the social issue dimension may be rare, it is nevertheless an important, and sometimes crucial, aspect of legislative and electoral politics, and it is difficult to answer some rather basic questions in the study of American politics without a multi-dimensional mapping of district-level preferences.

This calls for a survey-based approach to measuring preferences. In this study, we overcome the challenge of small sample sizes at the district-level by employing a novel Bayesian hierarchical model to estimate the median of voters' economic and social ideology at the district level. Our analysis combines two important methodological insights. First, we address the measurement error associated with individual survey questions by using a Bayesian Item-Response (IRT) model to generate a survey-based estimate of the median economic and moral preferences in each congressional district.[1] Unlike factor-score based approaches, our IRT model fully characterizes the measurement error associated with individual survey responses. Our estimates, however, are still subject to large sampling errors due to the relatively small sample sizes in each district. We overcome this problem by employing a Bayesian hierarchical model to "shrink" the estimates of district preferences toward the mean of the overall distribution conditional on a district's aggregate demographics. This approach enables us to generate estimate for the median issue preferences in each district that capture all of the information available from both survey responses and district-level demographic information. It also enables us to fully characterize the uncertainty in our estimates.

[1] See Clinton, Jackman, Rivers (2004), Treier and Jackman (2008), and Treier and Hillygus (2009) for useful introductions to IRT models.

Our approach to the estimation of district-level preferences makes a number of important contributions. First, our model enables researchers to generate estimates that fit the median voter models of political economy prevalent in the formal literature better than proxies based on vote share or demographics used in previous studies (Kernell 2009). For some of the most important unanswered questions introduced above, it is crucial to understand the distribution of policy preferences rather than partisanship or demographics across districts. Second, our approach illustrates the importance of considering multiple dimensions of ideology. Though we focus on economic and "moral values" issue dimensions, our approach could easily be extended to capture additional dimensions of voter ideology (e.g., foreign policy).

After explaining the methodology behind the estimates, we use them to provide some new insights into the distribution of preferences across U.S. Congressional districts, introducing some new stylized facts and providing further insight into some old ones. First, we underscore the importance of both the economic and moral values issue dimensions in understanding elections and representation in the United States, and we are able to identify the many districts where preferences on the economic and moral values dimensions are on opposite sides of the national median. Second, we demonstrate that the distribution of preferences across districts on the moral values dimension has a far wider, more platykurtic shape than that for the economic dimension, where most districts are tightly clustered around the national median. Third, we demonstrate that both distributions have a pronounced left skew akin to the distribution of Republican presidential votes across districts, indicating that the urban, Democrat-dominated districts in the left tail of the distribution are further from the national median than are the rural and exurban Republican districts on the right.

Finally, we provide a brief preview of the ways in which our estimates might help answer some of the questions introduced above. Given the unimodal inter-district distribution of preferences, our estimates throw cold water on the notion that gerrymandering or residential sorting explain the difference between the polarization of Congress and the apparent moderation of voters. Furthermore, our observation that there is a much wider distribution of district medians on the moral values dimension than the economic dimension might be a useful starting point for a theory about why party gatekeepers try so hard to avoid votes like that for the Stupak amendment, and keep bills related to moral values off the agenda, and hence out of the data sets used by Congressional scholars. Differences in the distributions on the two dimensions might also explain why preferences on the moral values dimension are better at predicting presidential than Congressional election results, and why some districts split their votes across these offices. Next, we argue that the left skew in the distribution of preferences across districts has potentially important implications for the literature on electoral bias and responsiveness.

In section 1, we discuss the approaches of previous studies. Next, we discuss the details of our empirical model. In section 3, we present and validate the empirical results. In section 4, we introduce stylized facts about the distribution of preferences across districts that can be gleaned from our estimates. In section 5, we discuss how these might inform ongoing debates in American politics, and section 6 concludes.

1. Previous literature on district preferences

There are five approaches that previous scholars have used to estimate ideology and issue preferences at the sub-national level.

1.1. Survey approaches

The most straightforward approach to estimating district preferences is to use data from a representative survey that asks respondents for their preferences on individual issues (Erikson, Wright, and Mclver 1993; Brace, Sims-Butler, Arceneaux, and Johnson 2002). The problem is that national surveys generally do not have enough respondents to develop efficient estimates of voter's preferences at sub-state levels. In their seminal study of legislative representation, Miller and Stokes used data from the 1958 American National Election Study to estimate policy preferences at the district level (Miller and Stokes 1963). However, their study had an average of just 13 respondents per district (see Achen 1978; Erikson 1978). While their estimates of constituency opinion were unbiased, they had extremely large standard errors.

Two recent studies have turned to the National Annenberg Election Survey (NAES) to overcome the Miller-Stokes problem. Clinton (2006) combines data on self-identified preferences from surveys conducted in 1999 and 2000 by Knowledge Networks (KN) and the NAES. However, although the two surveys have over 100,000 combined responses, congressional districts still have as few as 41 respondents. There also may be significant measurement error associated with self-identified ideology (Stiglitz 2009).

Instead of using self-identified ideology, Peress (2008) estimates a factor score of each respondent's ideology along several dimensions. He then takes the mean of these factor scores in each congressional district along the primary issue dimensions. While this approach is a major step forward, it does not characterize the uncertainty in the estimates of either individual-level ideology or the mean ideology in a district, which could lead to significant bias in subsequent analyses that relies on the estimates, especially given the small sample sizes in some districts. Another recent innovative study by Bafumi and Herron (2007) uses a Bayesian item response

model using a matrix of binary questions from the CCES that resembled roll-call votes to characterize the preferences of voters and their representatives in a common space.

1.2. Demographics

Another approach employed by a number of previous studies is to measure constituency policy preferences using demographic variables (e.g., Kalt and Zupan 1984; Peltzman 1984). This approach is often utilized in the congressional literature when researchers seek a proxy for district preferences. In this literature, researchers typically estimate a model in which legislative roll-call behavior is modeled as a function of district demographic characteristics obtained from the U.S. Census and other sources. However, estimating constituent preferences solely based on demographics requires several potentially problematic assumptions. First, demographic attributes are generally considered antecedents of ideology, rather than indicators of it. As a result, there is no guarantee that demographic characteristics alone enable us to locate congressional districts on a partisan or ideological continuum. Second, researchers must assume that individuals' demographic characteristics are related systematically to their policy preferences (Ardoin and Garard 2003), but demographic factors typically only explain a small amount of the variance in either individual or district-level preferences and vote choice.

1.3 Simulations

A related approach is to use demographic data to simulate district preferences. This methodology takes advantage of demographic data that are available at the district level, as well as knowledge concerning the relationship between individuals' demographic characteristics and their policy positions (Ardoin and Garand 2003). There are two types of simulations. In traditional simulations of constituency opinion, scholars utilize a lower level of aggregation to simulate opinion at a higher level of aggregation. For instance, they might use data from

individual-level surveys to simulate district-level preferences. Another approach is to use survey data from a higher level of aggregation to simulate opinion at a lower level of aggregation. For instance, they might use state-level data to simulate district-level preferences. However, both types of simulations are subject to unknown errors and employ strong parametric and measurement assumptions (Clinton 2006).

1.4. Electoral returns

Election returns are another popular way to estimate district preferences (e.g., Canes-Wrone, Cogan and Brady 2002; Ansolabehere, Snyder and Stewart 2001; Erickson and Wright 1980). The advantage of this approach is that it is explicitly based on electoral behavior and is thus linked to the partisan or ideological continuum that generally underlies electoral competition (Levendusky, Pope, and Jackman 2008). In addition, it is available across all states and districts and it is updated frequently (Kernell 2009).

However, there are a number of significant problems with using electoral returns to estimate preferences. First, presidential vote shares in any given election may be products of short-term forces rather than ideology (Levendusky, Pope, and Jackman 2008). For instance, LeoGrande and Jeydel (1997) find only moderate correlations for state-level presidential election results between most adjacent elections, which suggests that the reliability of the aggregate presidential vote is not extremely high. Second, even if short-term forces could be removed, the medians of district preferences can only be ranked ordinally based on presidential vote share if researching are willing to assume equal variance across districts (Kernell 2009). In order to address a variety of substantive questions of interest, for example about potential electoral and policy bias owing to the distribution of voters across districts, it is important to have district-level preference estimates that link up to a scalar conceptualization of an issue space. Third, it is

impossible to measure the preferences of sub-constituencies (e.g., the preferences of Democrats or Latinos) using presidential level vote shares (Clinton 2006). Finally, it is impossible to measure multiple issue dimensions using only electoral returns.

1.5 Bayesian Approaches

The most recent development in the literature is to use a Bayesian approach to measure district ideology. Park, Gelman, and Bafumi (2004) and Lax and Phillips (2009) apply Bayesian approaches to measure state ideology using multi-level regression and post-stratification (MRP). This approach combines survey-based data on individual issue preferences with individual-level geographic and demographic data to generate improved estimates of state-level ideology. Unfortunately, the MRP approach requires the Public Use Microdata Sample from the Census that is not available at the district-level.

Levendusky, Pope, and Jackman (2008) apply a Bayesian approach to estimate district partisanship by combining information about electoral returns with district demographics and a variety of other factors. They control for the effects of election-specific, short-term factors, such as national-level swings specific to particular elections, incumbency advantage, and home-state effects in presidential elections. This approach represents a novel improvement over previous attempts to measure district partisanship. But it has many of the same problems as other approaches that use election returns. First, the functional relationship of the mean of district vote shares and the median of district preferences is unclear (Kernell 2009). Second, it is impossible to measure the preferences of sub-constituencies using this method since presidential vote shares are only available at the district-level. Above all, it is impossible to disentangle different issue dimensions using this approach.

2: Model for District-level Preferences

Our approach builds upon many of the strengths of previous attempts to estimate district ideology. We use the relatively large samples within districts afforded by the 2004 Annenberg National Election Study to estimate individuals' latent preferences using a Bayesian IRT Model (Bafumi and Herron 2007, Treier and Hillygus 2009).[2] We aggregate these preferences to estimate the median in each district, using a Bayesian Hierarchical Model to address the problem of small sample sizes in many districts. This is, in essence, a marriage of the survey and demographic approaches, and borrows from the strengths of each, enabling us to draw on information from the entire distribution of district preferences to make inferences regarding the median preferences in each district. The advantage of our approach is that it enables us to incorporate multiple pieces of information to estimate both individual and district-level ideology. It also enables us to fully characterize the uncertainty at each stage of our analysis, as well incorporate this uncertainty into our final estimates of the posterior distribution of the median ideology in each district.

2.1. Characterizing Individual Issue Preferences

The simplest approach to estimating preferences would be to use survey respondents' self-identified preferences. Studies of this sort constitute a large body of work in elections and political behavior (e.g., Clinton 2006; Erikson, Wright, and McIver 1993; Fiorina 2004). However, recent work shows that people scale their self-identified ideology relative to a local distribution of preferences rather than the national distribution (Stiglitz 2009). In addition, self-identified preferences exhibit very little variation, making it problematic to use them to estimate the median in each district. Respondents overwhelmingly describe themselves as moderates, and Americans seem to have an aversion to describing themselves as liberals, even when their

[2] We use the Annenberg Election Survey rather than the CCES largely because we prefer the Annenberg policy questions to those in the CCES.

answers to policy questions suggest otherwise. Thus we evaluate voters' preferences using responses to survey questions on various issues.

Recent research by Shafer and Claggett (1995), Ansolabehere, Rodden, and Snyder (2006, 2008), Treier and Hillygus (2009), and Ensley (2008) suggests that answers to questions about economic issues are powerful predictors of individual vote choice, but there is also a second dimension related to moral and social values like abortion and gay rights that has gained salience since the 1980s. As a result, we estimate two dimensions, which aim to capture the economic and moral dimensions of voters' political ideology. The economic issue scale includes questions about pocketbook issues such as taxes, redistribution, healthcare, and union favorability, while the social issue dimension includes questions about issues such as abortion and gay rights.[3]

Previous researchers have used several strategies to map responses from policy questions to a scalar characterization of policy preferences or ideology. The simplest method is to simply sum responses to create one composite policy score (e.g., Heath, Evans, and Martin 1994). Another widely-used approach is to use factor analysis to create a weight for each policy question. The estimates of ideology through this approach consist of a linear combination of responses, with weights derived from factor analysis (e.g., Ansolabehere, Rodden and Snyder 2006; Bartels 2006). Finally, following the literature on ideal point estimation for legislators

[3] An alternative approach is to aggregate a wide assortment of policy questions and proceed by analyzing only the first dimension (Leblanc 2007), or the first two dimensions (Peress 2008) that arise from exploratory factor analysis. When we follow this approach, the first dimension appears to be dominated by a group of "social values" questions on which respondents' answers are highly correlated. Yet the scores generated from this approach are much worse than our economic scale at predicting respondents' partisanship and their survey responses on economic issues, and thus clearly miss an important economic dimension of partisan conflict—one that might be measured with greater error, but that nevertheless has a powerful impact on voting behavior.

(e.g., Jackman 2001; Clinton, Jackman and Rivers 2004) scholars have recently started to use Bayesian item response (IRT) models to create estimates of policy preferences (Treier and Hillygus 2009). We use the IRT approach in this paper.

There are a number of advantages to the IRT approach, two of which are most relevant for this paper. First, a problem with the NAES design is that many of the issue questions roll in and out of the survey questionnaire. As a result, many of the respondents have missing data for a large number of issue questions. The IRT method enables us to estimate this missing data "on the fly" as part of the estimation procedure for voters' latent ideology. This enables us to incorporate uncertainty in the imputation of the missing data directly into our estimates of voters' ideology. Second, the IRT method treats voters' ideology as a "latent variable" that is measured with error. As a result, these traits are subject to inference and it is straightforward to calculate the uncertainty associated with our estimates of voters' latent ideology (Treier and Hillygus 2009). We can then take the uncertainty in our estimates of individual voters' ideology into account for our estimates of district-level ideology.

2.2 Data.

Our approach relies upon 21 questions from the rolling cross-section component of the 2004 NAES. This survey was conducted between October 2003 and November 2004 in the lower 48 states and the District of Columbia. The NAES employs random digit dialing to generate its sample, which yields representative samples at both the national and sub-national levels. The survey also yielded an unusually large number of respondents, with N = 81,322. A disadvantage of the 2004 NAES, however, is that many of the questions were only asked for a relatively short span of time. Moreover, many of the questions were asked to a random sub-sample of survey respondents. This creates a substantial missing data problem. Fortunately, however, the IRT

approach enables us to impute missing data as part of the estimation procedure. This enables us to include respondents that have answered as few as 8 questions, including at least 4 questions along each dimension, in our analysis. This reduces the total number of respondents we consider in our analysis to N=31,802.

2.3 The IRT Model of Individual Issue Preferences

We estimate individual issue preferences using the Bayesian Item-Response (IRT) model described in Treier and Jackman (2008) and Treier and Hillygus (2009). We model individual issue responses as a function of the unobserved preference dimension via an ordinal item-response model. Let $i = 1, \ldots, n$ index individuals and $j = 1, \ldots, m$ index issues. Let $k = 1 \ldots K_j$ index the (ordered) response categories for issue j; in our data, K_j ranges from two to five.

Then our model is

$$\begin{aligned}
&\Pr(y_{ij} = 1) = F(\tau_{j1} - u_{ij}) \\
&\Pr(y_{ij} = 2) = F(\tau_{j2} - u_{ij}) - F(\tau_{j1} - u_{ij}) \\
&\vdots \\
&\Pr(y_{ij} = k) = F(\tau_{jk} - u_{ij}) - F(\tau_{j(k-1)} - u_{ij}) \\
&\vdots \\
&\Pr(y_{ij} = K_j) = 1 - F(\tau_{j,(K_j-1)} - u_{ij})
\end{aligned}$$

where $u_{ij} = x_{1i}B_{1j} + x_{2i}B_{2j}$, y_{ij} is the i-th respondent's answer to question j, x_i is the two-dimensional latent ideology score for respondent i, and $F(\cdot)$ is the logistic CDF, $F(z) = \frac{\exp(z)}{1+\exp(z)}$. The slope parameters B_{1j} and B_{2j} are the item discrimination parameters, tapping the extent to which variation in the dth latent dimension generates different response probabilities.

The unknown parameters in our IRT model are $\theta = \{\underset{(nxd)}{x}, \underset{(dxm)}{\beta}, \tau\}$ where $\tau = (\tau_1', \ldots, \tau_m')$. The likelihood for the data is:

$$L(\theta \mid Y) = \prod_{i=1}^{n} \prod_{j=1}^{m} \Pr(y_{ij}; x_i; \beta_j; \tau_j)$$

We work in a Bayesian setting to explore the joint posterior density of the model parameters using a Markov chain Monte Carlo (MCMC) method (see Jackman 2000, 2004, 2009, and Gelman and Hill 2007 for useful overviews of MCMC models). We implement this MCMC scheme using the statistical packages R and RJags (Plummer 2009).[4] We use diffuse normal priors for the discrimination parameters B_j, with mean zero and variance 100. The first cutpoint is assigned the same diffuse normal prior as B_j. The subsequent cutpoints are assigned exponential priors with mean 1/2. We specify normal priors with mean 0 and variance 1 for each x_{di}.

In order to ensure rotational invariance and fix the scale of the model, we assign fixed discrimination parameters to two reference items.[5] We also fix a number of the survey questions to only load on our dimension to make our analysis conceptually consistent with previous research on voters' ideology (Ansolabehere, Rodden, and Snyder 2006, 2008; Treier and Hillygus 2009).[6] After estimation, we post-process our results to normalize each dimension of latent preferences to have a mean of 0 and a variance of 1. We also post-process the discrimination parameters in our IRT model to ensure they remain on the same scale as our estimates of latent preferences (see Treier and Jackman 2008).

[4] The start values for x_{di} are created from an additive index, for each dimension, of all indicators that load on that dimension (scaled by the number of questions answered by the respondent for each dimension). We let the algorithm run for 2,000 iterations as burn-in, moving away from the start values such that subsequent iterations represent samples from the joint posterior density. Estimates and inferences are based on two chains of 10,000 iterations each, thinned by 40, in order to produce 500 approximately independent draws from the posterior density.

[5] Based on a preliminary logit analysis, we fix the discrimination parameters on the economic dimension for question ccc41 regarding the role of the government in addressing inequality to 1.5 and the discrimination parameters on the moral dimension for the question cce01 regarding abortion to -2.

[6] See Appendix A for more information about the questions.

2.4. Bayesian Hierarchical Model of District-level Preferences

Next, we use our estimates of individual preferences along each latent dimension to estimate the median economic and moral preferences in each congressional district.[7] We examine medians rather than means in order to be consistent with the large body of work in policy economy that draws inferences based on the median voter in a geographic district (e.g., Downs 1957). While this approach generates relatively unbiased estimates of the median preferences for each district, many of the estimates have large sampling errors due to the generally small sample sizes in each district. Indeed, the average congressional district contains just 79 respondents along each dimension. We use a bootstrapping procedure to estimate the standard errors of the median preferences in each district.[8] This approach incorporates both the uncertainty in our estimates of individual-level latent ideology, as well as the sampling errors stemming from the relatively small number of respondents in each district.

The bootstrapping approach reveals that there is not enough information in any single district to estimate its median preferences with any great precision. Fortunately, we have information from 431 other districts.[9] If we assume that these districts are exchangeable conditional on some set of covariates, we can use a Bayesian hierarchical model to draw on information from the entire distribution of district preferences to make inferences regarding the median preferences in each district (Jackman 2009, 301).

Specifically, we assume that the preferences in each district can be modeled as a function

[7] We estimate the medians by taking the median in each district at each iteration of our MCMC output, and then taking the median across iterations.
[8] The bootstrapping procedure generates 100 samples of the respondents in each district at each of the 500 iterations of the MCMC posterior distribution. Next, it calculates the median from each sample. Finally, it estimates the standard deviation of the distribution of that statistic.
[9] The NAES does not include Alaska and Hawaii.

of its district-level demographics (i.e., they are exchangeable conditional on these covariates).[10] Using a weighted least squares model, we generate a predicted value for the median economic and moral ideology in each district. We then partially pool these predicted values with the observed preferences to an extent determined by the relative uncertainty in the observed medians and the predicted medians. The final estimates of each district's preferences are thus an optimal combination of the information from that district and the relevant information in other districts.

More formally, we can write our model as:

$$(1)\ y_i \sim N(\theta_{i,}, \sigma_i^2)$$

$$(2)\ \theta_{i,} \sim N(\mu_i, \omega_i^2)$$

$$(3)\ \mu_{i,} = X_i^{\copyright} B$$

$$(4)\ \omega_i^2 = \tau^2 \times \sigma_i^2$$

Each observed district median y_i based on our Bayesian IRT estimates is a draw from a normal density with unknown mean $\theta_{i,}$ and variance σ_i^2. We calculate σ_i^2 using the bootstrapped estimate of the sample variance for y_i. The estimates of $\theta_{i,}$ (the preferences in district$_i$ that we would observe given an arbitrarily large number of survey respondents) are drawn from a normal distribution with an unknown mean μ_i and unknown variance ω_i^2. We regard each district as conditionally exchangeable. Thus, we can model μ_i as a function of district-level covariates $X_i^{\copyright}$. We include a variety of district-level demographic variables that have been found by previous individual-level studies to be correlated with economic or moral ideology.

Since the variances in each district vary, we obtain ω_i^2 by multiplying τ^2 by the variance of each district's preferences σ_i^2 (which is the equivalent of weighting each observation by

[10] The district-level demographics include percent urban, percent black, percent Latino, and median household income (each taken from the U.S. Census), as well as percent veteran, average church attendance, and the percentage belonging to unions (each taken from NAES).

$1/\sigma_i^2$). This generates a weighted least squares estimator for μ_i with the districts with smaller standard errors given larger weights and districts with larger standard errors given smaller weights (Gelman and Hill 2007, 389). The degree of pooling (the "shrinkage factor") between μ_i and y_i for each district depends on the variance σ_i^2 of our observed values for each district and the variance ω_i^2 from our weighted least squares model (see Gelman and Hill 2007, 477). In order to complete the specification of our Bayesian hierarchical model, we specify vague, non-informative values for the hyperparameters, letting the data dominate inferences.

We estimate the model through Markov-Chain Monte Carlo (MCMC) algorithms using the statistical packages R and RJags. We initialize the sampler with a random draw from the prior densities defined above, and then let the sampler run for 25,000 iterations using four parallel chains. We then discard the first 5,000 iterations as burn-in, and sample every 80th remaining iteration to estimate our posterior densities.

3. Results: Measuring Congressional District Ideology

In this section, we present the results of our multi-step Bayesian model of district issue preferences along the economic and social/moral dimensions. First, we present the results of our individual-level Bayesian IRT model of voter preferences. Next, we present the results of our Bayesian hierarchical model of district preferences. Third, we discuss several validation analyses of our results. Finally, we discuss the relationship between district-level demographic factors and district issue preferences.

3.1 IRT Results

The latent estimates of individual issue preferences run from a liberal extreme of -2.1 to the conservative extreme of 2.99 on the economic dimension, and liberal extreme of -2.37 to the

conservative extreme of 2.21 on the social dimension, with a mean of 0 on both dimensions.[11] The two scales are only modestly correlated (.27), indicating that at least two issue dimensions are necessary to account for voters' issue preferences. Our measures of economic and social ideology are correlated at .37 and .5 with self-reported ideology, and they are correlated at .47 and .39 with a standard 7-point party identification scale. These are similar to the correlations found by Ansolabehere, Rodden, and Snyder (2008) and Treier and Hillygus (2009).

3.2 Characterizing District Preferences

Each congressional district's median economic and social ideology appears in our model as a parameter to be estimated. Using the Markov Chain Monte Carlo algorithm, we obtain many samples from the joint posterior density. We then use these posterior densities to estimate the median estimate and 95% credible intervals for the economic and social preferences in each district, as well as the effect of each demographic covariate on district preferences.

FIGURE 1 HERE

Figure 1 displays point estimates for the median of each district's latent ideology along the economic and social dimensions; the thin gray lines are 95% credible intervals for each district. The graph indicates that the uncertainty associated with our estimates of district ideology is high relative to the differences across districts. Indeed, our posterior estimates of individual districts' economic ideology have an average standard deviation of .05, while the between-district standard deviation in our estimates of district economic ideology is 0.13. The relatively large uncertainty in the estimates for individual districts makes it difficult to make fine distinctions among districts. Fortunately, however, our Bayesian approach enables us to fully characterize the uncertainty in our estimates for the latent economic and social ideology of each district. It

[11] Appendix A presents more details about the discrimination parameters from our IRT model.

also enables us to directly compare any two districts to estimate the probability that one district is more liberal than another. For instance, our best guess is that the economic ideologies of the CA-07, represented by Democrat George Miller and the CA-08, represented by Democrat Nancy Pelosi are -0.10 and -0.22, respectively. The probability that the CA-08 is more liberal than the CA-07 along the economic dimension is 93%.

3.3. Validating the District Ideology Measures

We can further assess the validity of our new measures of district economic and social preferences by seeing whether they predict relevant political outcomes better than the mean self-reported ideology in the NAES, a measure that has been used in past empirical work to estimate district preferences (e.g., Clinton 2006). We evaluate the performance of each measure by regressing the results from each iteration of our model output for district preferences against 1) the Democratic vote share in the 2004 presidential election and 2) the Democratic vote share in the 2004 congressional election.[12] This approach enables us to fully evaluate the potential effect of measurement error associated with our estimates of latent district preferences on subsequent estimates of the effect of ideology on other political outcomes. We report the median estimates and 95% credible intervals of our model outputs in Table 1.[13]

TABLE 1 HERE

Combined, our two dimensions of preferences perform at least as well as, and generally much better than, the mean self-reported ideology from the NAES in predicting all three political outcomes. Most notably, over 70% of the variation in Democratic vote shares in the 2004

[12] We use data for congressional elections collected by Gary Jacobson. In uncontested elections, the winning candidate is assigned 75% of the vote.

[13] The credible intervals reflect the 2.5 and 97.5 percentiles of the aggregated results from running Bayesian regressions on the output of each of the 500 iterations of our Bayesian hierarchical model of district preferences.

presidential election is explained by our Bayesian estimates of district preferences, while the means of the self-reported ideology in each district explain just 62% of the variation. The Bayesian estimates of issue preferences also explain more of the variation in Congressional vote shares and DW-Nominate scores in 2004 than self-reported ideology.

Interestingly, the vast majority of the variation in presidential and congressional vote shares, as well as DW-Nominate scores, is explained by a district's economic rather than moral preferences—an observation that is consistent with a growing skepticism about the centrality of the "culture war" in American politics (Ansolabehere, Rodden, and Snyder 2006, Bartels 2006). However, we also see that the moral values dimension clearly has an impact on the electoral behavior of districts, one that comes through most clearly in presidential as opposed to Congressional elections. Moreover, it is somewhat surprising to see that district-level moral values preferences have an impact on Congressional voting on the first dimension of DW-Nominate, which is widely understood to be dominated by votes on bills related to economic policy. We also find that House members' votes on salient social-issue roll calls are highly correlated with their districts' moral value preferences. For instance, we find that the economic dimension of district preferences is uncorrelated with Democratic representatives' votes on the recent Stupak amendment to the healthcare bill, while the moral dimension is strongly predictive of Democrats' votes.[14] We explore the possible implications of these observations below.

3.4 Relationship of Demographic Factors with District Preferences

[14] The Stupak amendment is roll call # 884 in the 111th Congress. We have also used the Bayesian IRT methodology to estimate ideal points for representatives based on a subset of "social" roll calls identified by the Family Research Council, National Right to Life Committee, NARAL, and Human Rights Campaign. In a model that controls for party, we find that district-level social preferences are a much better predictor of social roll-call votes than are economic preferences. Moreover, as with the Stupak amendment, this effect is driven disproportionately by Democrats.

What is the relationship of demographic factors with latent district economic and social ideology? Figure 2 presents parameter estimates of the relationship between various demographic characteristics and district ideology. It shows the median and 95% credible intervals for each variable.

FIGURE 2 HERE

For the most part, the relationships we find between district ideology and demographic characteristics would not surprise many scholars of American politics. First, districts with high proportions of African Americans are consistently among the most liberal districts. Second, districts with greater church attendance and a higher fraction of the population that are evangelicals are significantly more conservative along both dimensions. A district's religiosity is a particularly strong predictor of its moral ideology. Indeed, our model indicates that church attendance and the percentage of a district's population that is evangelical have twice as large an effect on a district's social ideology as any other variable. Evangelicals constitute more than 60% of the population in each of the ten most conservative districts along the social dimension, while they constitute less than 16% of the population in the ten most liberal districts. Third, districts with more veterans are more conservative along both dimensions. Fourth, districts with greater union membership are significantly more liberal along the economic dimension, but more conservative along the moral dimension. Similar to Gelman et al. (2008) and De la O and Rodden (2008), we find that richer districts (as measured by median income) are substantially more conservative along the economic dimension, but we find no relationship between a district's median income and its median ideology on the moral dimension. Finally, we find that more urban districts more liberal on the moral dimension.

4. The Geographic Distribution of Preferences across U.S. Congressional Districts

4.1. Mapping Preferences

One of the benefits of our two-dimensional approach is that we can capture the political geography of preferences in a way that is sensible given our outside knowledge of the world.

FIGURE 3 HERE

In Figure 3, we use cloropleth maps to display the geography of both dimensions of district preferences. We find that New England congressional districts are generally liberal along both dimensions. However, our results also support the conventional wisdom that New Hampshire is more liberal on social issues than economic ones. Indeed, we find that the NH-01 in eastern New Hampshire is quite liberal along the social dimension. It is relatively conservative, however, along the economic dimensions. For instance, to examine some representative survey questions, 55% of respondents in the NH-01 are in favor of reducing income inequality compared to over 59% nationwide. Moving south, we find that many congressional districts in western Pennsylvania are extremely liberal along the economic dimension, but quite conservative on the social dimension. For instance, the PA-12, represented by Democrat John Murtha, is more liberal than 74 percent of other districts on the economic dimension but more conservative than 73 percent of districts on the social dimension.

The maps also illustrate the complex political geography of the Southeastern United States. Indeed, many poor, rural congressional districts in the south are relatively liberal along the economic dimension, but these same districts are among the most conservative in the nation along the moral dimension. For instance, 56% of the respondents in Arkansas's rural, strongly religious 1st District favor banning all abortions, while 59% favor policies that aim to reduce income differences. Much of the Midwest is also relatively liberal on the economic dimension, but much more conservative on social issues. In the west, the opposite pattern holds, with many

districts much more liberal on the social dimension than the economic dimension. For instance, in some wealthy suburban districts like the California 48th in Orange County and the Washington 8th in the Seattle suburbs, voters are extremely liberal along the social dimension, with around 20 percent favoring a ban on abortions, but more conservative on the economic dimension, with less than 50 percent in favor of reducing income inequality.

4.2. Issue Preferences and Congressional Elections

Overall, we stress that a significant advantage of our model is that it enables us to understand the two-dimensional dynamics of American elections much better than simpler one-dimensional models.

FIGURE 4A HERE

A useful way to visualize preferences in two dimensions is with Figure 4a, in which the economic estimates are on the horizontal axis and social preferences are on the vertical axis, with national medians indicated by the dashed lines. Solid red and blue dots are districts won consistently by Republicans and Democrats, respectively, in each of the 2004, 2006, and 2008 elections. The hollow blue dots switched from Republican to Democrat control during this period, and the hollow red dots experienced the opposite switch.

Figure 4a demonstrates that the two dimensions are correlated, but only weakly so, and well over one third of the districts are to the left of the median district on one dimension but to the right on the other. The Figure also drives home the observation that the partisanship of districts is explained in part by both dimensions, but the economic dimension is far more powerful. Visualizing district preferences in two dimensions also helps clarify the nature and extent to which the partisan affiliation of members of Congress is "out of step" with the policy preferences of their constituents. Many of these districts follow the example of Murtha's

Pennsylvania 12th, where the preferences match the party position on one dimension and not the other. In such "split" districts, the more comfortable partisan fit is usually on the economic dimension. That is, the upper left quadrant is mostly blue, and the lower right quadrant is mostly red.

Yet there are also a number of districts in which the ideological fit is poor on both dimensions. The hollow blue dots in the lower left quadrant indicate that the Republicans controlled a handful of such districts, but lost most of them in the Democratic tidal waves of 2006 and 2008. We also see a rather large cluster of blue dots in the upper right quadrant, where districts that are center-right on economics and rather far to the right on moral values are represented by Democrats. Most of these are in the South, and many have adopted the "blue dog" moniker. In accordance with popular perception, the Democrats have a large and potentially unwieldy coalition on both dimensions, but they have been able to achieve victory by being especially flexible on the social dimension, where their members of Congress represent the entire gamut from the extreme right to the extreme left.

Consistent with the literature on incumbency bias, we also observe that the vast majority of "out of step" districts are controlled by longstanding incumbents, such as Republican Ros-Lehtinen in the liberal Florida 18th, and on the other side of the aisle, Vic Snyder in the Arkansas 2nd and Chet Edwards in the Texas 17th.

4.3. The Distribution of Preferences across Districts

Another advantage of our approach is that it allows for an examination of the shape of the distribution of preferences across districts. Figure 4b presents kernel densities of the estimates for both dimensions.

FIGURE 4B HERE

Two things stand out in Figure 4b. First, both distributions demonstrate a pronounced left skew resembling that of the district partisanship estimates generated by Levendusky, Pope, and Jackman (2008), and the distribution of Republican presidential votes across Congressional districts. On both issue dimensions, the districts on the far left are substantially further from the median district than are the districts on the far right. On both dimensions, the left tail is made up almost exclusively of densely populated urban districts that vote overwhelmingly for Democrats at all levels of government—the lower-left hand corner in Figure 4a. The skew is particularly pronounced for the economic dimension. While the individual-level IRT estimates of economic preferences display a modest right skew, individuals are distributed across districts such that left-wing voters are more geographically concentrated than right-wing voters. This striking stylized fact has a number of potentially important implications that will be discussed below.

Second, Figure 4b demonstrates a much wider, platykurtic distribution of median preferences on the moral values dimension when contrasted with the tight, leptokurtic distribution of preferences on the economic dimension. The standard deviation across districts on the estimated moral values median is almost twice as large as that on the economic dimension. Except for the urban districts in the left tail, the median voter in most districts is quite moderate on the economic dimension. In contrast, there is a much smaller density of districts around the median on the moral values dimension, and a larger number of relatively extreme districts. However, unlike the well-known bimodal distribution of DW-Nominate scores, both distributions are clearly unimodal.

5. Implications

While interesting in their own right, these stylized facts have implications for several inter-related literatures in American politics.

5.1. Candidate Positioning

Congressional scholars have used surveys of candidates, combined with information from roll-call votes of those who have served in Congress, to measure candidate positions (Ansolabehere, Snyder, and Stewart 2001; Burden 2004). A key finding in these studies is that for both parties, there is a positive correlation between the district-level presidential vote and the positions of candidates, which indicates that candidates are able to tailor their platforms to the ideology of their districts. Yet there is virtually always a gap between the position of the Democratic and Republican candidate—one that is larger in the more ideologically extreme districts—which indicates that parties clearly place strong constraints on the abilities of candidates to converge all the way to the district medians. This constraint might come from the importance of primary constituencies (Gerber and Morton 1998), or it might be the case that candidates simply cannot credibly adopt platforms that vary substantially from that of the party leadership, since voters know that once in office, the party leaders will use a variety of tools to whip co-partisans on important votes, and voters can observe that in practice, party-line voting is commonplace on roll-call votes.

These papers also establish some additional empirical regularities without explaining them. First, there is a greater spread of party positions across Democratic than Republican candidates. Second, the correlation between district preferences (captured by presidential voting) and candidate positions is stronger for Democratic than Republican candidates. Third, the gap between candidates is larger in the far-left districts than in the far-right districts.

Each of these curiosities makes good sense in light of our estimates. The pronounced left skew in district median preferences means that there are more districts just to the right of the national median than just to the left, and the right tail is much closer to the median district than

the left tail. Whatever it is that makes it difficult for candidates to buck their party's platform—the primary constituency or a credibility problem—presumably constrains both parties equally as one moves away from the median district in either direction. But there is a rather large group of urban districts in the left tail that is much further from the median district than is the most right-wing district, where Republican candidates simply cannot adopt platforms that would make them competitive. Democrats can compete in a larger number of "ideologically Republican" districts than vice-versa. As we can see clearly in Figure 4 above, this also translates into more "out of step" victories for the Democrats.

It is not surprising, then, that we see a more diverse pool of candidates among the Democrats, and a more homogeneous one among the Republicans.[15] Moreover, it stands to reason that the correlation between district-level presidential vote and candidate platforms will be higher for the Democrats, since credible Republicans are simply not competing in the districts of the left tail. This might also explain why the gap between candidates is larger in the far-left districts than in the far-right districts.

However, our estimates also call for two notes of caution about this literature. First, the use of roll-call voting to simulate candidate positions (Fiorina 1974, Ansolabehere, Snyder, and Stewart 2001) might lead to exaggerations in the extent to which candidates appear to diverge. Second, the use of presidential votes as a proxy for district preferences might be flawed in these analyses if presidential elections are driven by different issues than are Congressional elections. We now address each of these points in turn.

5.2. Differences between Congressional Ideal Points and Public Preferences

Fiorina (2004) sparked a lively debate by pointing out that while scales based on roll-call

[15] For a similar observation, see Grofman, Merrill, Brunell, and Koetzle (1999) and Mayer (1996).

voting indicate a bimodal distribution of Congressional ideal points in recent years, the distribution of preferences across voters appears to be unimodal during the same period, and most voters appear to have moderate preferences on both economic and moral issues. A related puzzle is that while it appears that at least two dimensions are needed to capture the preferences of the electorate (Poole 2002; Layman and Carsey 2006), roll-call voting in Congress is captured by a single dimension (Poole and Rosenthal 1997).

One possible explanation of the difference in polarization is that while there is a unimodal distribution of individuals, they are sorted into homogeneous Congressional districts, either by their own location decisions or by incumbent-friendly redistricting, such that a bimodal distribution of district medians emerges. The unimodal kernel densities in Figure 4b call for skepticism (see also Peress 2008; McCarty, Poole, and Rosenthal 2009). Moreover, it is puzzling that the spread of median preferences across districts on the moral values dimension is far wider than that on the economic dimension, while it is the latter that has driven the increasing partisan divergence and polarization of roll-call voting (McCarty, Poole, and Rosenthal 2006).

Perhaps our two-dimensional estimates can serve as a starting point for new theoretical approaches to these puzzles. One possibility is that due to the problem addressed in Snyder (1992), scales based on roll-call votes generate artificial polarization because of the non-random process through which roll-call votes are called. Fiorina (2009) points out that strategic party gatekeepers might be less inclined to propose changes from the status quo that require roll-call votes among issues on which the party is internally divided. Divergence across districts is greater on the moral values dimension (Figure 4b), and in practice, the parties in Congress are more internally divided on this issue dimension (Figure 4a), with the majority party often owing its control of the agenda to the fact that it has managed to capture a good number of districts on

the opposite side of the national median. Thus when looking for bills to champion, majority-party leaders might shy away from moral values issues that would expose the party's fissures and endanger their most vulnerable incumbents, leaving any changes from the status quo on these issues to the courts, state legislatures, or executive orders. Rather, if there are costs to twisting the arms of moderate co-partisans, party gatekeepers might be more inclined to introduce bills that change the status quo when this would be attractive to the median voter in most of the districts controlled by the party. Given the intra-district distributions we have presented, more often than not, these will be on economic issues. This would help explain the dominance of economic issues in Nominate scores and the surprising paucity of votes on moral values issues. Moreover, such strategic agenda control might help explain why the partisan polarization of such scores is so pronounced relative to the preferences of voters.

Perhaps a more detailed study of candidate platforms, if allowing for multiple dimensions, might reveal that the lack of roll-call votes on moral values issues allows candidates from the two parties to converge to the district median, whereas since representatives' feet will regularly be held to the fire by roll-call votes on economic issues, their party labels prevent them from credibly converging to the district median on this dimension.

5.3. Differences between Congressional and Presidential Elections

We demonstrated above that the moral values dimension does a better job explaining vote choice in presidential than in Congressional elections, and both parties are able to win in districts with moral preferences that are quite dissonant with median preferences. This makes sense if moral values issues are kept off the Congressional agenda, giving the candidates more leeway to converge to the district median on the moral than the economic dimension. But in presidential elections, the candidates of the two parties are forced to adopt a single national platform on

moral values issues, and the president clearly holds important powers to issue executive orders and nominate Supreme Court justices.

This perspective might help explain patterns of split-ticket voting. In districts where the moral values dimension is highly salient, voters might choose a member of Congress that adopts a moral values platform that is close to the district median but out of step with her party, while choosing the presidential candidate with the most proximate position on the moral values dimension. For example, a morally conservative but economically moderate district might choose an anti-abortion Democrat while voting for the Republican presidential candidate. Indeed, preliminary analysis of the 2004 election suggests that in the districts that split their Congressional and presidential votes, the moral values dimension is uncorrelated with Congressional votes, but it is an extremely powerful predictor of the presidential vote.

The left skew of the distribution of preferences across districts might also help shed light on an additional puzzle about split-ticket voting: over the course of the post-war period, the vast majority of split districts have been won by Republican presidential candidates and Democratic Congressional candidates rather than the other way around. The geography of preferences might help explain this. If we assume that the presidential candidates' platforms are symmetrically arranged around the national median in the long run, it is potentially important that there are more districts just to the right of the national median than just to the left, and that the right tail is relatively close to the national median. This implies that more "Republican" districts are winnable for ideologically flexible Democrats than vice-versa.

5.4. Electoral Bias and Responsiveness

Our preference estimates suggest that left-wing individuals are more concentrated within (largely urban) electoral districts than are right-wing individuals. This might help explain recent

empirical analysis suggesting that once incumbency bias and uncompetitive districts are properly accounted for, long-term electoral bias in Congressional elections has been consistently pro-Republican. Quite simply, Democrats are more concentrated in space than Republicans. However, the left skew in the distribution of preferences across districts may have a silver lining for Democratic candidates. We have shown that there is a relatively large density of center-right districts in which they are competitive, and this might allow them to experience disproportionately large gains in conservative districts when benefiting from valence shocks related to war, economic conditions, or scandals.

From the perspective of normative Democratic theory, a more interesting question is whether the distribution of preferences across districts might produce distortions in the transformation of preferences to public policies. Again, our estimates encourage closer analysis. On the economic dimension, because of the concentration of leftists in urban districts, the median voter in our sample of individuals is well to the left of our estimate of the preference of the median voter in the median district, raising the possibility that under plausible assumptions about the policy process, the profile of policies produced by Congress will be to the right of what one might expect under a simple application of the median voter theorum that ignored geography.

6. Conclusions

The contributions of this paper are both methodological and substantive. First, we have developed a simple approach that enables scholars to combine survey data with widely available census demographic data to estimate district preferences in a variety of contexts. We have illustrated how this approach can be used to estimate congressional district preferences, but there are a variety of other potential applications. For instance, our approach could be used to

estimate the preferences of state legislative districts, counties, cities, or virtually any other sub-state political unit. It would be straightforward for scholars to augment our model with additional demographic data. Our approach provides a rigorous and flexible framework for combining multiple sources of information to estimate district ideology along multiple dimensions.

The second contribution is a useful new description of American political geography that arises from our estimates. Above all, in spite of the focus of Congressional scholars on a single economic dimension, we stress the importance of understanding U.S. elections and representation in two dimensions. Three stylized facts stand out. First, a large number of districts are to the right on one dimension but to the left on the other, and while both dimensions help explain election outcomes, the economic dimension is more powerful, especially in Congressional elections. Second, the spread of median preferences across districts is much wider on the moral than on the economic dimension. Third, the distributions of both issue dimensions demonstrate a left skew across districts that mimic the distribution of presidential votes across districts.

Finally, we have suggested ways in which these stylized facts might inform some ongoing theoretical and empirical debates in American politics. We attempted to illuminate some of the ways in which an understanding of the geography of preferences in two dimensions might help scholars develop new approaches to some classic questions about the nexus that links the policy preferences of voters, the platforms of candidates, and the voting behavior of representatives to the policies that are produced. Previously, the unavailability of district-level issue preferences has forced scholars to use a variety of second-best proxies for preferences. Our results enable scholars to directly examine how cross-sectional variation in district ideology

might affect the platforms candidates choose in the districts, and the issues that party leaders bring to the floor of the legislature. Understanding the two-dimensional distribution of preferences across districts might allow for new insights into what appear to be puzzling differences between the distributions of Congressional and voter ideal points, and between the outcomes of Congressional and presidential elections. Moreover, our estimates have potential to help explain electoral bias and responsiveness, and perhaps even some distortions in the transformation of preferences to policies.

References

Achen, Christopher. 1978. ‘‘Measuring Representation.’’ *American Journal of Political Science* 22 (3): 475–510.

Ansolabehere, Stephen, James M. Snyder Jr. and Charles Stewart III. 2001. "Candidate Positioning in U.S. House Elections." *American Journal of Political Science* 45: 17–34.

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

Ansolabehere, Stephen, Jonathan Rodden, James M. Snyder, Jr. 2008. "The Strength of Issues: Using Multiple Measures to Gauge Preference Stability, Ideological Constraint, and Issue Voting." *American Political Science Review* 102: 215-232.

Ardoin, Phillip and James C. Garand. 2003. Measuring Constituency Ideology in U.S. House Districts: A Top-Down Simulation Approach. *Journal of Politics.* 65(4): 1165-1189.

Bafumi, Joseph and Michael Herron. 2007. "Preference Aggregation, Representation, and Elected American Political Institutions." Paper presented at the 2007 Annual Meeting of the Midwest Political Science Association, Chicago, IL.

Bartels, Larry. 2006. "What's the Matter with What's the Matter with Kansas?" *Quarterly Journal of Political Science* 1: 201-226.

Brace, Paul, K Sims-Butler, K Arceneaux, M Johnson. 2002. Public Opinion in the American States: New Perspectives Using National Survey Data. *American Journal of Political Science*. 46(1): 173-189.

Burden, Barry. 2004. "Candidate Positioning in U.S. Congressional Elections." *British Journal*

of Political Science 34: 211-227.

Canes-Wrone, Brandice, John F. Cogan and David W. Brady. 2002. "Out of Step, Out of Office: Electoral Accountability and House Members' Voting." *American Political Science Review*. 96(1): 127:140.

Clinton, Joshua. 2006. ''Representation in Congress: Constituents and Roll Calls in the 106th House.'' *Journal of Politics*. 68(2): 397-409.

Clinton, Joshua, Simon Jackman, and Doug Rivers. 2004. "The Most Liberal Senator? Analyzing and Interpreting Congressional Roll Calls." *PS: Political Science and Politics* 37: 805-11.

De La O, Anna and Jonathan Rodden. 2008. "Does Religion Distract the Poor? Income and issue Voting around the World." *Comparative Political Studies* 41(4-5): 437-76.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper Collins.

Ensley, Michael. 2008. "Core Values, Value Conflict, and Ideological Identification." Working paper, Kent State University, Department of Political Science.

Erikson, Robert S. 1978. ''Constituency Opinion and Congressional Behaviour: A Reexamination of the Miller-Stokes Representation Data.'' *American Journal of Political Scienc*e 22 (3): 511–35.

Erickson, Robert and Gerald C. Wright. 1980. "Policy Representation of Constituency Interests." *Political Behavior* 2: 91–106.

Erikson, Robert S., Gerald C. Wright, Jr. and John P. McIver. 1993. *Statehouse Democracy: Public Opinion and the American States*. Cambridge University Press.

Fiorina, Morris. 1974. *Representatives, Roll-Calls, and Constituencies*. Lexington, MA: Lexington.

Fiorina, Morris. 2004. Culture War? The Myth of a Polarized America. New York: Pearson Longman.

Fiorina, Morris and Samuel Abrams. 2009. *Disconnect: The Breakdown of Representation in American Politics*. Norman, OK: University of Oklahoma Press.

Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York and Cambridge: Cambridge University Press.

Gelman, Andrew, David Park, Boris Shor, Joseph Bafumi, and Jeronimo Cortina. 2008. *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way they Do*. Princeton, NJ: Princeton University Press.

Gerber, Elizabeth and Rebecca Morton. 1998. "Primary Election Systems and Representation," *Journal of Law, Economics, and Organization* 13: 304-324.

Grofman, Bernard, Samuel Merrill, Thomas Brunell and William Koetzle. 1999. "The Potential Electoral Disadvantages of a Catch-All Party," *Party Politics* 5: 199-210.

Heath, Anthony, Geoffrey Evans, and Jean Martin. 1994. "The Measurement of Core Beliefs and Value." *British Journal of Political Science* 24: 115-158.

Jackman, Simon. 2000. "Estimation and Inference via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo." *American Journal of Political Science* 44 (2): 375–404.

Jackman, Simon. 2001. Multidimensional Analysis of Roll Call Data via Bayesian Simulation: Identification, Estimation, Inference, and Model Checking. Political Analysis 9(3): 227-241.

Jackman, Simon. 2004. "Bayesian Analysis for Political Research." *Annual Review of Political Science* 7: 483–505.

Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. West Sussux, U.K.: John Wiley & Sons Ltd.

Kalt, Joseph P. and Mark A. Zupan. 1984. "Capture and Ideology in the Economic Theory of Politics.'' *American Economic Review* 74: 279-300.

Kastellec, Jonathan P. and Eduardo L. Leoni. 2007. "Using Graphs Instead of Tables in Political Science." *Perspectives on Politics*. 5(4): 755-771.

Kernell, Georgia. 2009. "Giving Order to Districts: Estimating Voter Distributions with National Election Returns." *Political Analysis* 17(3): 215-235.

Lax, Jeffrey and Justin Phillips. 2009. "Gay Rights in the States: Public Opinion and Policy Responsiveness." *American Political Science Review* 103: 367-86.

Layman, Geoffrey and Thomas Carsey. 2006. "Changing Sides or Changing Minds? Party Identification and Policy Preferences in the American Electorate." *American Journal of Political Science* 50: 464-477.

Leblanc, William. 2007. "Party Positions and the Seats/Votes Relationship with Ideological Voters." Dissertation, Department of Political Science, MIT.

LeoGrande, William, and Alana S. Jeydel. 1997. "Using Presidential Election Returns to Measure Constituency Ideology: A Research Note." *American Politics Quarterly* 25(1): 3-19.

Levendusky, Matthew S., Jeremy C. Pope and Simon Jackman. 2008. "Measuring District Level Preferences with Implications for the Analysis of U.S. Elections." *Journal of Politics* 70: 736-53.

Mayer, William. 1996. *The Divided Democrats: Ideological Unity, Party Reform, and Presidential Elections*. Boulder, CO: Westview Press.

McCarty, Nolan, Keith Poole, and Howard Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge, MA: MIT Press.

McCarty, Nolan, Keith Poole, and Howard Rosenthal. 2009. "Does Gerrymandering Cause Polarization?" *American Journal of Political Science* 53(3): 666-680.

Miller, Warren E. and Donald E. Stokes. 1963. ''Constituency Influence in Congress.'' *American Political Science Review* 57: 45-56.

Park, David, Andrew Gelman, and Joseph Bafumi. 2004. "Bayesian Multilevel Estimation with Poststratification: State-Level Estimates from National Polls." *Political Analysis* 12(4): 375-85.

Peltzman, Sam. 1984. ''Constituent Interest and Congressional Voting.'' *Journal of Law and Economics*. 27: 181--210.

Peress, Michael. 2008. "Candidate Positioning and Responsiveness to Constituent Opinion in the U.S. House of Representatives." Unpublished paper, Rochester University.

Plummer, Martyn. 2009. JAGS: Just Another Gibbs Sampler. Version 1.0.3. http://calvin.iarc.fr/~martyn/software/jags/

Poole, Keith. 2002. "Recovering a Basic Space from a Set of Issue Scales." *American Journal of Political Science* 42: 954-993.

Poole, Keith and Howard Rosenthal. 1997. *Congress: A Political Economic History of Roll Call Voting*. New York: Oxford University Press.

Shafer, Byron E., and William JM Claggett. 1995. *The Two Majorities: The Issue Context of American Politics.* Johns Hopkins University Press.

Snyder, James M. 1992. "Artificial Extremism in Interest Group Ratings," *Legislative Studies Quarterly*, Vol. XVII, No. 3 (August), pages 319-345.

Stiglitz, Edward H. 2009. "Liberal Relative to Whom? Ideology and Policy Preferences in the American Public." Working Paper.

Treier, Shawn and Sunshine Hillygus. 2009. "The Nature of Political Ideology in the Contemporary Electorate." *Public Opinion Quarterly*. 73(4): 679–703.

Treier, Shawn and Simon Jackman. 2008. "Democracy as a Latent Variable." A*merican Journal of Political Science*. 52(1): 201–217.

Table 1: Validation of Results

Table 1a: Relationship between District Preferences and Presidential Election Results

(All coefficients standardized)	Logit Model of 2004 District-Level Democratic Presidential Outcomes		OLS Model of 2004 District-Level Democratic Presidential vote shares	
	(1)	(2)	(3)	(4)
Intercept	37.97 [30.51, 46.37]	-1.10 [-1.52, -.69]	7.51 [7.04, 7.98]	-.11 [-.16, -.05]
Mean Self-Reported Ideology	-12.12 [-14.79, -9.75]		-2.36 [-2.53, -2.10]	
Bayesian Estimate of Economic Pref.		-17.17 [-22.33, -12.76]		-3.46 [-3.98, -2.98]
Bayesian Estimate of Social Pref.		-5.79 [-7.99, -3.85]		-.77 [-1.04, -.51]
R-squared	.50 [.49, .51]	.51 [.49, .54]	.61 [.60, .61]	.71 [.68, .74]
95% credible intervals in brackets				

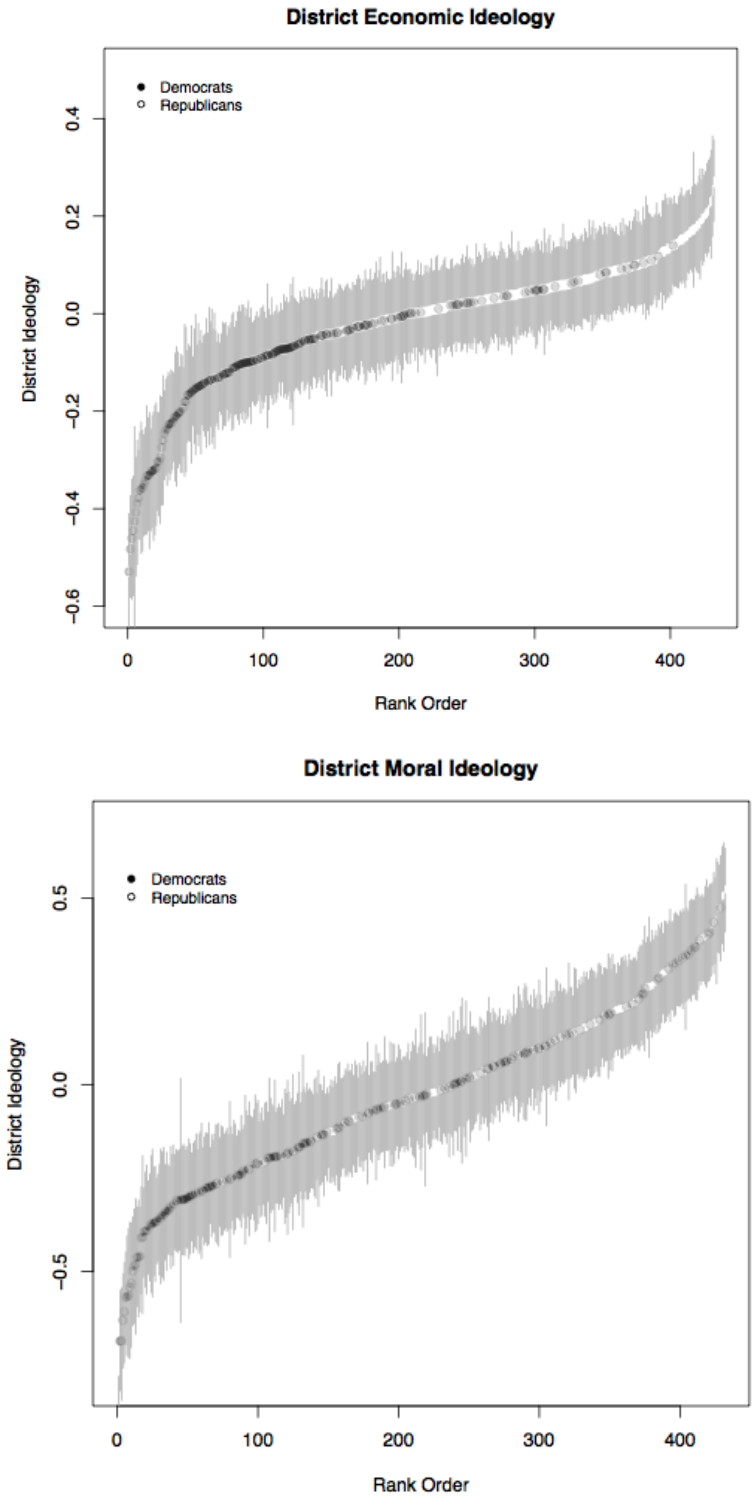
Table 1b: Relationship between District Preferences and Congressional Election Results

(All coefficients standardized)	Logit Model of 2004 District-Level Democratic House Outcomes		OLS Model of 2004 District-Level Democratic House vote shares	
	(1)	(2)	(3)	(4)
Intercept	20.69 [17.2, 25.5]	-.41 [-.72, -.10]	8.33 [7.44, 9.26]	-.08 [-.16, .00]
Mean Self-Reported Ideology	-6.56 [-8.0, -5.5]		-2.61 [-2.90, -2.33]	
Bayesian Estimate of Economic Pref.		-15.17 [-19.5, -11.5]		-4.29 -5.03, -3.60]
Bayesian Estimate of Social Pref.		-.37 [-1.85, 1.17]		-.46 [-.85, -.06]
R-squared	.27 [.26, .28]	.34 [.32, .37]	.41 [.41, .41]	.48 [.46, .50]
95% credible intervals in brackets				

Table 1c: Relationship between District Preferences and Congressional Voting Behavior

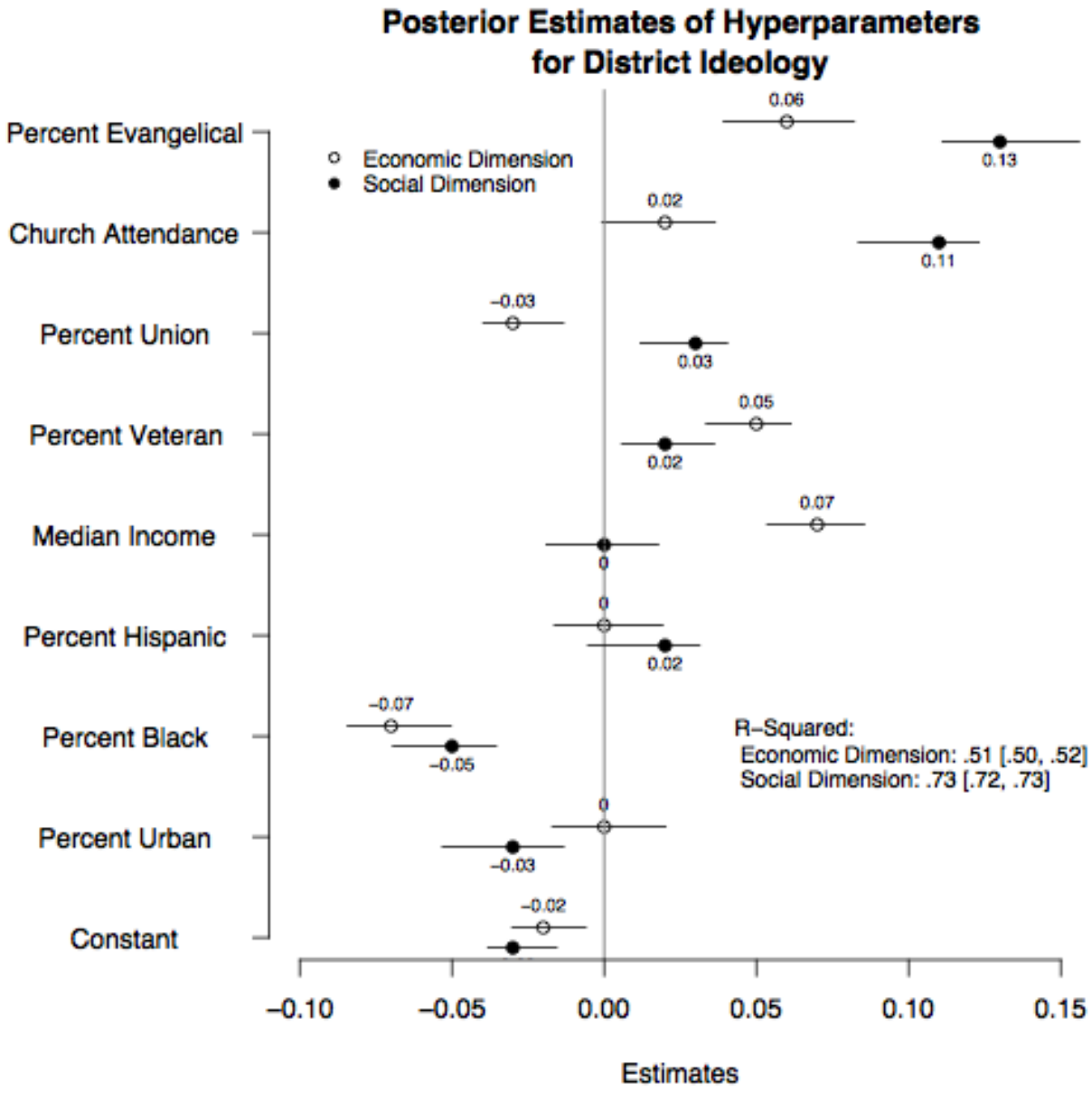
(All coefficients standardized)	OLS Model of Nominate Scores (109th Congress)		Logit Model of Votes on Stupak Amendment to Healthcare Bill (112th Congress)	
	(1) (All Members)	(2) (All Members)	(3) (All Members)	(4) (Only Democrats)
Intercept	-4.34 [-4.70, -3.83]	.14 [.09, .18]	0.77 [0.46 , 1.11]	-0.80 [-1.26, -0.35]
Mean Self-Reported Ideology	1.39 [1.23, 1.50]			
Bayesian Estimate of Economic Pref.		2.22 [1.82, 2.65]	7.45 [4.61, 10.72]	0.134 [-3.90, 4.00]
Bayesian Estimate of Social Pref.		.30 [.07, .52]	6.03 [4.37, 7.88]	8.50 [6.09, 11.43]
R-squared	.37 [.36, .37]	.41 [.39, .42]	0.38 [0.36, 0.40]	0.39 [0.35, 0.43]

Figure 1: District Ideology
Medians and 95% Credible Intervals



These figures show the rank order, medians, and 95% Credible Intervals for our estimates of district economic and moral ideology.

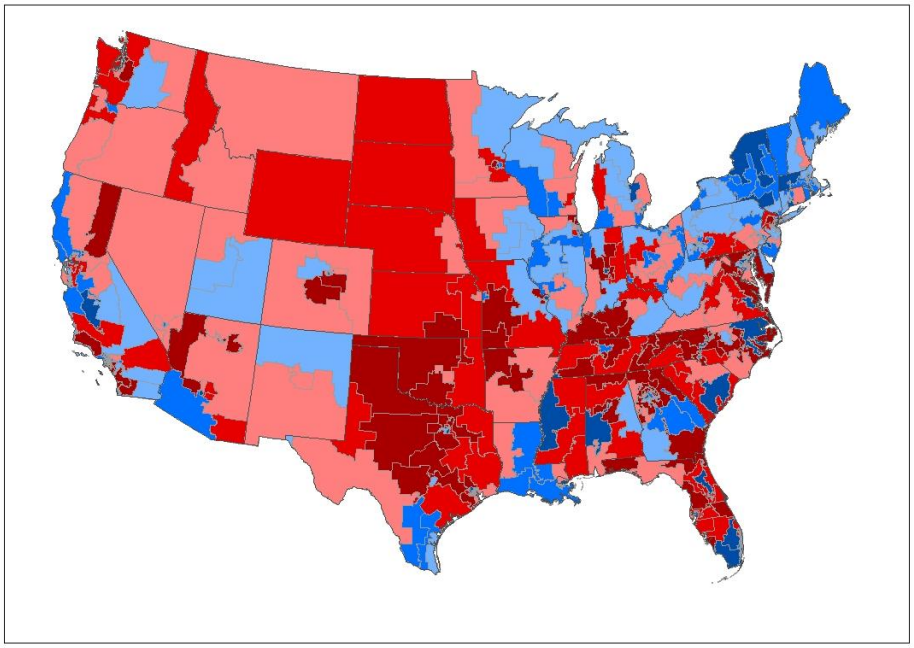
Figure 2



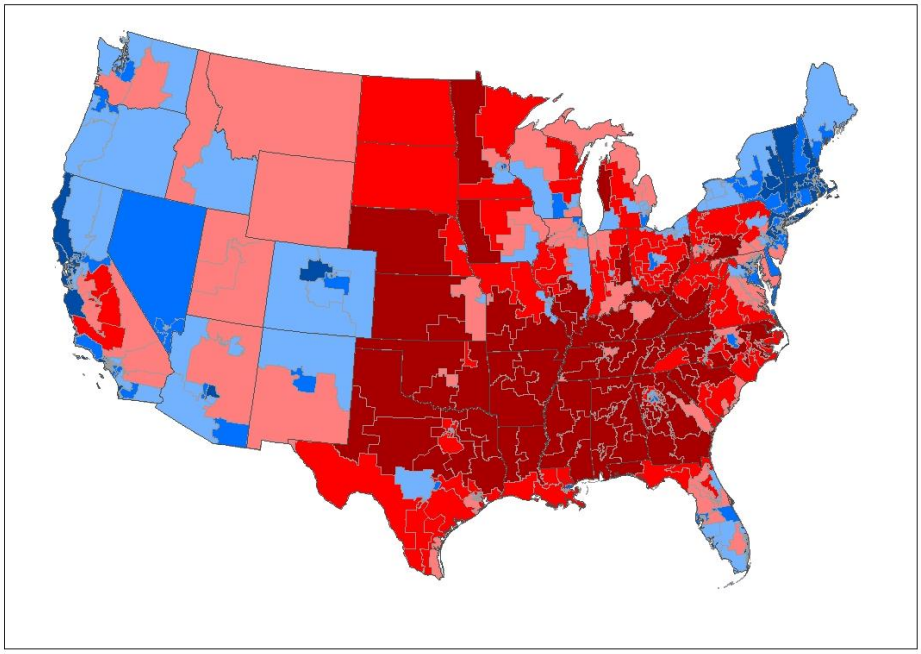
This figure shows the relationship between standardized district demographic characteristics and latent district economic and social/moral ideology. It shows the mean of our estimate for each hyperparamter as well as the 95% credible interval.

Figure 3
Geographic Distribution of Congressional District Preferences

Congressional District Economic Ideology



Congressional District Moral Ideology



Bluer areas are more liberal and redder areas are more conservative.

Figure 4a: Economic and Social Estimates by Congressional Party

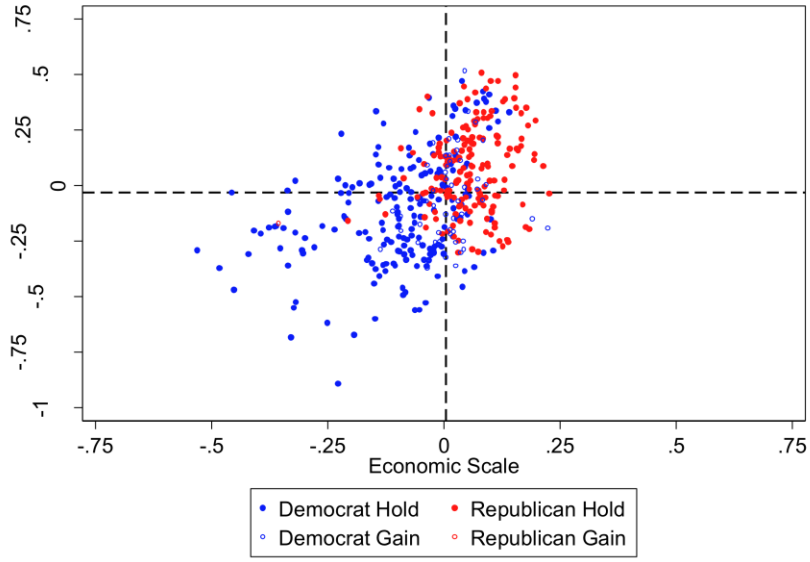
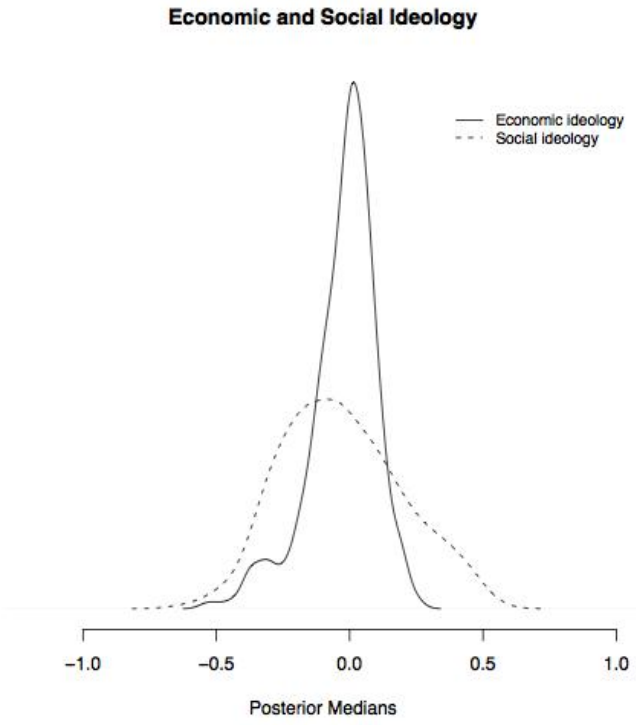


Figure 4b:



This figure plots the density of our estimates of district economic and moral ideology.

Appendix A

Table A.1: Questions Used in Factor Analysis for Construction of Issue Scales

	Economic Dimension		Social Dimension	
	Discrimination Parameter	**95% HPD Interval**	**Discrimination Parameter**	**95% HPD Interval**
Health insurance for children	-2.41	[-3.08, -1.91]	0	
Health insurance for workers	-1.22	[-1.490, -0.969]	0	
Eliminating estate tax	0.500	[.337, .677]	0	
Investing social security in stock market	0.417	[.335, .502]	0	
Making labor organizing easier	-1.200	[-1.540, -0.924]	0	
Labor union favorability	-1.15	[-.124, -1.05]	0	
Federal funding for schools	1.67	[1.52, 1.85]	0	
Government should reduce inequality	1.5		0	
Health insurance for everyone	2.05	[1.85, 2.26]	0	
Reduce taxes	-0.318	[-0.383, -0.256]	0	
School vouchers	-0.1200	[-.1830, -0.0584]	-0.547	[-0.601, -0.491]
Immigration	-0.344	[-0.443, -0.254]	-0.334	[-0.419, -0.244]
Gun Control	0.818	[0.726, 0.918]	0.286	[0.212, 0.371]
Feminist groups favorability	-1.020	[-1.130, -0.919]	-1.13	[-1.23, -1.05]
Abortion	0		-2	
Gay groups favorability	0		-2.14	[-2.33, -1.99]
Christian groups favorability	0		0.893	[0.812, 0.970]
Gay Marriage (federal)	0		-1.54	[-1.64, -1.46]
Gay Marriage (states)	0		2.47	[2.18, 2.79]
Late Term Abortion	0		-1.070	[-1.170, -0.959]
Stem cells	0		-1.34	[-1.61, -1.09]