Appendix: Measuring Aggregate-Level Ideological Heterogeneity

Matthew S. Levendusky* and Jeremy C. Pope†

July 13, 2009

*Assistant Professor, Department of Political Science, University of Pennsylvania. Email: mleven@sas.upenn.edu

†Assistant Professor, Department of Political Science and Research Fellow, Center for the Study of Elections and Democracy, Brigham Young University. Email: jpope@byu.edu

1 Data

We begin by describing the factor analysis we used to initially explore the data and assess the plausibility of using items on particular dimension. We go on to give the exact question wording from the CCES for each item and the estimated item difficulty/discrimination parameters.

Factor Analysis and Item Selection

To select the items used in our analyses of the economic and social dimensions, we began by factor analyzing all of the "ideological" items included in the common content module of the CCES,[1] which results in 15 items for analysis. An eigenvalue decomposition suggests a 3-factor solution, which is represented in table 1.

The loadings in table 1 suggest that the first dimension is primarily an economic dimension of ideology, the second dimension is primarily a social dimension, and the third dimension is a "globalization" dimension dominated by attitudes towards trade and immigration.

We take the items from the first dimension and use those items to measure economic ideology, with a few exceptions. First, we exclude the items measuring attitudes towards Iraq and affirmative action. Iraq is not strictly speaking an "economic" attitude. And, though it has an economic component, affirmative action has strong racial and social components.[2] Second, we also exclude the union influence measure, since it was only asked of a small segment of the sample. That leaves us with the six issues discussed in the paper. Note that we did not include the liberal-conservative self-identification item in our factor analysis here. If we were to include it, it would load primarily on this economic dimension. We re-estimated

[1] Here, "ideological" items are any items that seem to tap a respondent's underlying ideological preferences, defined fairly broadly.

[2] It is beyond the scope of our paper to determine why these attitudes load on this economic dimension.

Variable	F1	F2	F3
Gay Marriage	-0.54	**0.62**	0.22
Abortion	-0.22	**0.84**	
Affirmative Action	**0.56**	-0.21	-0.26
Stem Cell Research	0.34	**-0.61**	
Partial-Birth Abortion	-0.48	**0.61**	
Illegal Immigration	-0.38	0.18	**0.57**
Union Influence	**0.63**	-0.24	
Social Security	**-0.69**	0.34	
Minimum Wage Vote	**-0.61**	0.24	
Capital Gains Tax Cut	**0.69**	-0.32	
Tax vs. Spending	**0.55**	-0.26	-0.16
Jobs vs. the Environment	**0.53**	-0.21	-0.11
CAFTA Roll-Call	0.28		**0.43**
Right to Invade Iraq	**0.78**	-0.46	
Withdraw Troops	**0.72**	-0.36	
Proportion of Variance Explained	0.31	0.18	0.05

Table 1: Exploratory factor analysis (3 factor solution) of the ideological items included in the common content module of the CCES. The largest factor loading for each item is given in **bold**.

our model for the economic dimension including the liberal-conservative self-identification item, and the results are highly consistent with the results given in the body of the paper. We excluded it on theoretical grounds (we wished to give a purely operational measure of economic ideology), but again, should scholars wish to see it included, they could easily replicate our procedure and produce a new set of estimates.

Taken together, these six economic items strongly suggest a one-dimensional structure. If we take the correlation matrix of these items, the first eigenvalue is 3.4, and the second eigenvalue is 0.61; examining a scree-plot of the eigenvalues supports the same conclusion of unidimensionality. Given this, we conclude that these items can profitably be analyzed as tapping attitudes along the same dimension of public opinion.

Specific Item Wordings: Economic Dimension

As we discussed in the paper, we used six items from the CCES to measure attitudes on the left-right economic dimension. In this section, we give the question wordings and response options for each of these items.

Five-Point Scale Items:

- **Social Security:** Now, we'd like to ask you about Social Security. A proposal has been made that would allow people to put a portion of their Social Security payroll taxes into personal retirement accounts that would be invested in private stocks and bonds. Do you favor or oppose this idea? [Response options are a five-point scaling ranging from strongly agree to strongly disagree]

- **Environment vs. Jobs:** Some people think it is important to protect the environment even if it costs some jobs or otherwise reduces our standard of living. Other people think that protecting the environment is not as important as maintaining jobs and our standard of living. Which is closer to the way you feel, or haven't you thought much about this? [Response options are a five point scale: (1) "Much more important to protect the environment even if lose jobs" (2) "Environment somewhat more important", (3) "About the same", (4) "Economy somewhat more important", and (5) "Much more important to protect jobs, even if environment worsens"]

Continuous Items:

- **Increase Taxes or Spending Cuts:** If your state were to have a budget deficit this year it would have to raise taxes on income or sales or cut spending, such as on education, health care, welfare, and road construction. What would you prefer more raising taxes or cutting spending? Choose a point along the scale from 100% tax increases (and no spending cuts) to 100% spending cuts (and 0% no tax increases).

The point in the middle means that any the budget should be balanced with equal amounts of spending cuts and tax increases. [Response options are a 0-100 scale as described in the text]

Binary Items:

- **Cut Domestic Spending First:** What would you most prefer that Congress do - cut domestic spending, cut military spending, raise taxes, or borrow funds? [Respondents select one of the four options. We code respondents who select "cut domestic spending" as 1, all others as 0.]
- **Minimum Wage Roll-Call Vote:** What do you think? If you were faced with this decision, would you vote for or against increasing the minimum wage? [Respondents can vote for or against raising the minimum wage, we code respondents who say they would vote for the increase as 1, those who vote against it as 0]
- **Capital Gains Tax Cut Roll-Call Vote:** What do you think? If you were faced with this decision, would you vote for or against these tax cuts? [Respondents can vote for or against extending the capital gains tax cut, we code respondents who say they would vote for the extension as 1, those who vote against it as 0]

2 Results: Average State Ideology, Economic Dimension

Which states are (on average) the most conservative? The most liberal? In figure 1, we give the posterior means and 95% highest posterior density intervals for the state mean parameter (μ_m, where m indexes states).

As figure 1 reveals, there is both a considerable amount of heterogeneity and homogeneity among the states. Heterogeneity in that various states are more or less conservative:

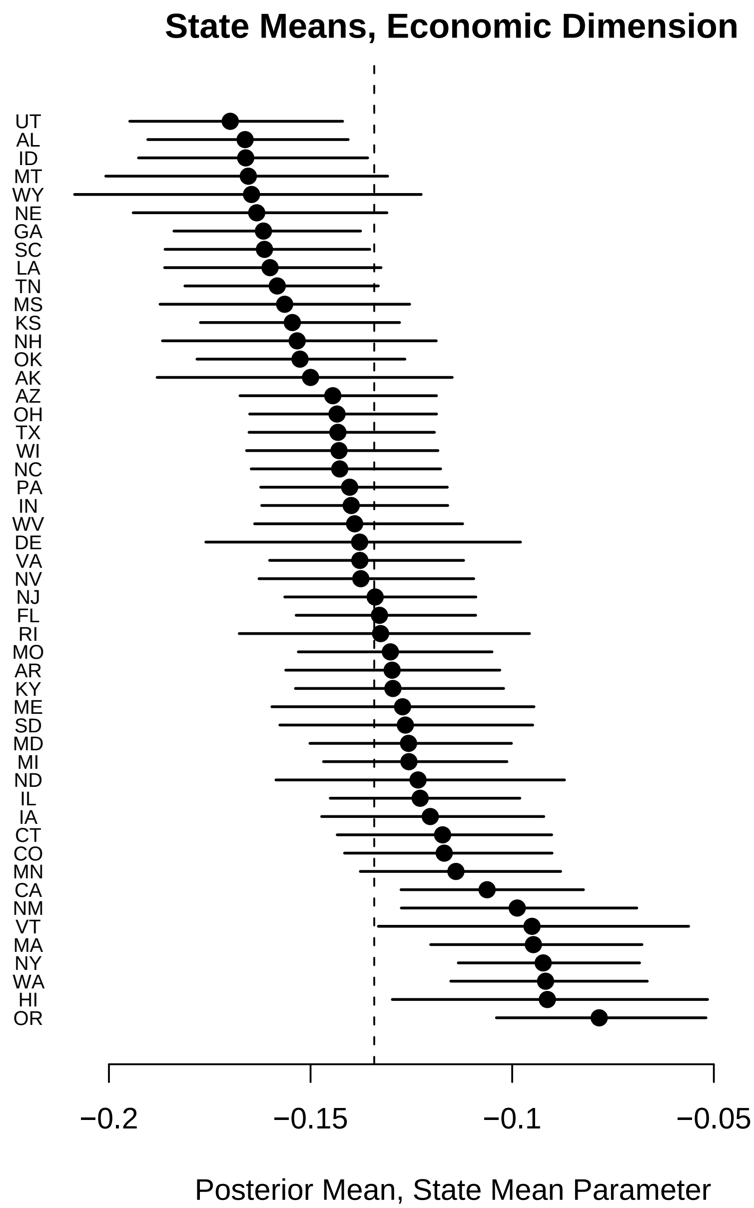

Figure 1: Posterior means and 95% highest posterior density intervals for the state mean parameters.

some states are quite liberal (OR, NY, MA—largely states on the Pacific coast and in New England), others more moderate (PA, FL, NJ) and others more conservative (UT, AL, ID–largely states from the Mountain West or the deep South). Based on our model, we would estimate Oregon, Hawaii, or Vermont to be the most liberal state in the union, with a posterior probability of 0.57, 0.21, and 0.14, respectively (the remainder of the posterior probability is split between a number of other states, each with a small posterior probability of being the most liberal in the nation). Likewise, we estimate Wyoming, Montana, Utah, or Idaho to be the most conservative state in the nation (probability 0.23, 0.17, 0.17, and 0.10, respectively). All of this largely conforms to the conventional wisdom about the geographic distribution of ideology in the US—the more liberal areas are clustered on the coasts and the more conservative spots are in the interior.

There are a few states that might seem a bit out of place. For example, our results indicate that New Mexico is estimated to be the among the 10 most liberal states. We wish to stress, however, that this is a function of the data and not our method. Even if we just took simple averages of the raw CCES data, we would estimate NM to be the 4th most liberal state, and using only the average liberal-conservative self-placement item, it would be the 6th most liberal state. In other words, the *data* on New Mexico indicate it is particularly liberal, not our scaling method. There are two possibilities. It is possible that New Mexicans are not really that liberal, but rather the CCES sample may not reflect the true distribution of opinion in New Mexico (particularly given the small sample of only 321 respondents). Alternatively, New Mexico's reputation for moderation may have more to do with a relatively socially conservative environment (a dimension not captured in the questions we selected). We leave exploring these possibilities for future research, simply noting here that these types of discrepancies are a function of the data, not the model.

Figure 1 highlights an important advantage of our method over many previous approaches—our model allows us to calculate the uncertainty in our estimates of the average ideology within each state. For example, looking at figure 1, we see a few clusters of states: a small group of very liberal states, a small group of very conservative states, and a large group of states in between.[3] In other words, once we take measurement error into account, states are less easily differentiated along this dimension than we might initially suspect.

To reinforce this point, we can compare the ideology of various states and ask how confident we can be that one state is more liberal than another. We consider two comparisons: Oregon vs. Washington (two liberal states), and Ohio and Pennsylvania (two moderate states). When comparing liberal or conservative states in the tails of the ideological distribution, we can usually make relatively fine distinctions: the probability that Oregon is more liberal than Washington is 0.92. But near the center of the distribution, making these sorts of distinctions becomes more difficult: the probability that Ohio is more conservative than Pennsylvania is 0.69, which falls far short of conventional levels of statistical significance. Once we take seriously the proposition that state ideology is not known with certainty, differentiating states becomes quite difficult.

This point has profound substantive consequences. It may seem simple to suggest that we have a limited ability to draw definitive conclusions about the relative ideology of states. It is anything but. Even in 2006—by many accounts a highly polarized age—ordinary Americans remain more heterogeneous at the state level than has generally been appreciated. That is, even states that one is tempted to think of as liberal or conservative bastions—say California or Texas—are in fact quite diverse. California has a number of conservative regions in the inland farming communities, and Texas has more liberal enclaves in Austin and other large cities like Houston. There are real limits to our ability to distinguish between the average voter in different states.

There is one small caveat to the above discussion. As with all statistical procedures, our

[3]Throughout the paper, when we say "state X is more liberal (conservative) than state Y," the reader should take this to mean "the mean of the voter distribution in state X is estimated to be more liberal (conservative) than the mean of the corresponding voter distribution for state Y." We adopt this rhetorical convention for simplicity.

model performs better with more data. This is hardly a revelation, but it has an important implication for our discussion about our ability to distinguish states from one another. In particular, the model can make more precise estimates of the mean for states with more data, as figure 2 reveals.

Figure 2 implies that we can more easily distinguish differences some states (with more data) than others. So there is one caveat that we need to apply to the discussion above: with more data (say, 3000 interviews per state), we could more accurately estimate the mean opinion within a given state–that is, the standard error of the mean is smaller in CA than it is in WY. But we wish to emphasize, however, that this is a secondary concern to the one we pointed out above: the primary limitation here is that most states are relatively similar places, at least in terms of average ideology (see the clustering of states in figure 1). Here, the small sample size is an issue, but it is a secondary issue. So the limitation is not so much in the sampling, but in the states themselves: because they're not all that different on average, it is difficult to distinguish states from one another.

3 Results: Social Issues Dimension

In the body of the paper, we outlined a method for estimating ideological heterogeneity at the U.S. state level along the left-right economic dimension. Here, to demonstrate the flexibility of our model, we use a slightly simpler model to generate estimates on the social issues dimension. We say "slightly simpler" because the model here is the same as the model described in the body of the paper, except that the random effects are only stratified by state, rather than by state and party (i.e., the model used in the district level analysis).

We select the items used to measure ideology along this dimension using the factor analysis described above, see table 1. This gives the following items: (1) the respondent's

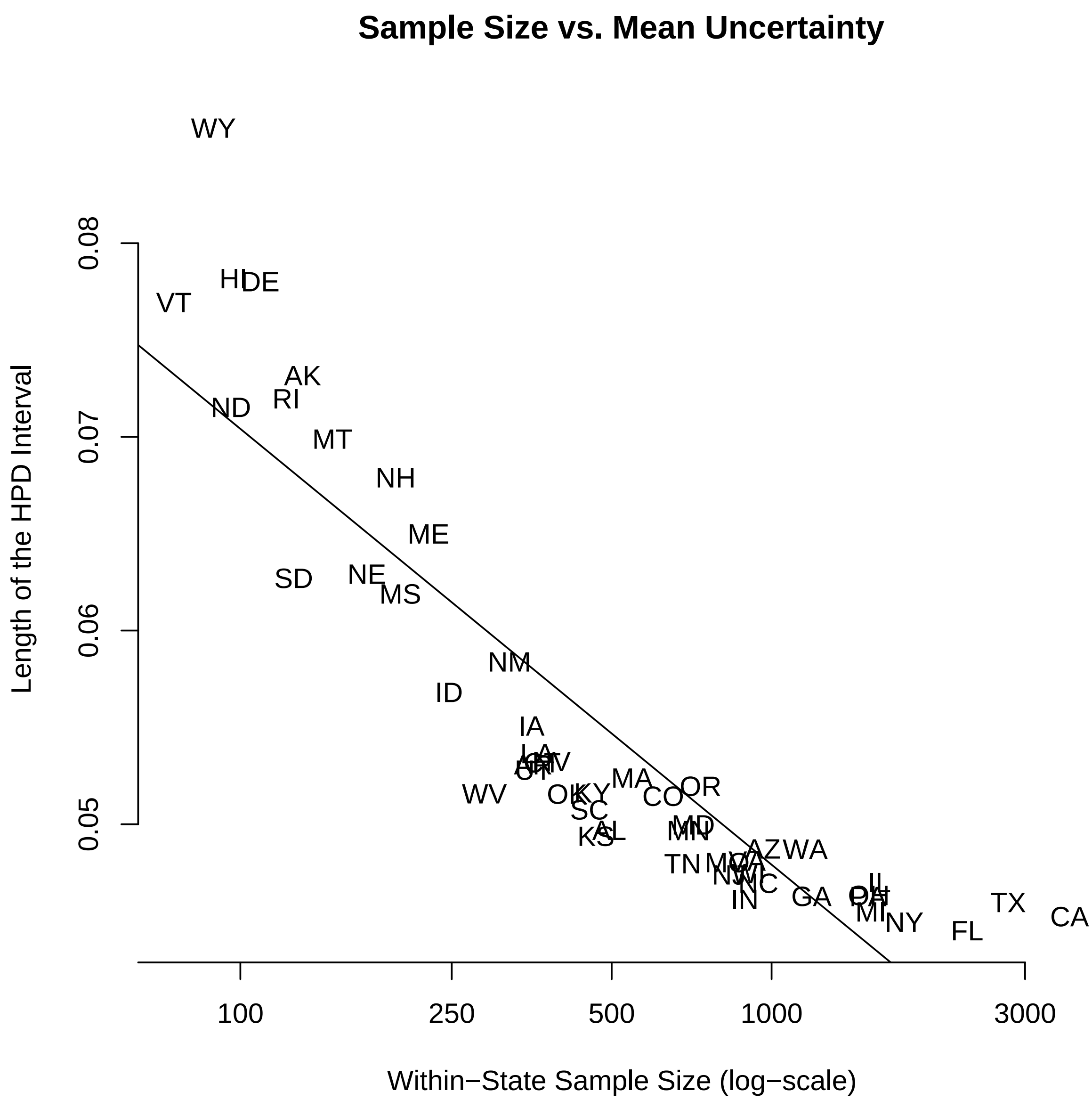

Figure 2: Plot of the length of the 95% HPD intervals for the state mean parameter (as estimate of the precision with which we can estimate the mean) against the (logged) within-state sample size. The solid line gives the OLS regression with all 51 states, the dotted line omits the estimate for Washington, D.C. (a clear outlier). The graph reveals a clear pattern, with states with more data estimated more accurately.

attitude on abortion,[4] (2) the respondent's attitude toward partial-birth abortion,[5] (3) the respondent's attitude toward stem cell research,[6] and (4) the respondent's attitude towards gay marriage.[7] Dimensional analysis suggests that these items are strongly one-dimensional as well: the first eigenvalue is 2.8, the second is 0.5. In light of this, we analyze these items as one-dimensional.

Before proceeding to the results, we should note an important caveat about our estimates on the social dimension. Here, the raw data fits less well with our underlying model, suggesting that respondents may be more polarized along this dimension. As such, our estimates here are more dependent on our model assumptions, and may be less robust to other modeling choices. Given this, scholars should exercise somewhat more caution when using these estimates.

[4]This item reads: "There has been some discussion about abortion during recent years. Which one of the opinions on this page best agrees with your view on this issue? (1) By law, abortion should never be permitted; (2) The law should permit abortion only in case of rape, incest or when the woman's life is in danger; (3) The law should permit abortion for reasons other than rape, incest or danger to the woman's life, but only after the need for the abortion has been clearly established. (4) By law, a woman should always be able to obtain an abortion as a matter of personal choice"

[5]This item was one of the "roll-call" items. Respondents were first given a brief description of the roll-call vote on the partial-birth abortion bill, and were then asked "How about you? If you were faced with this decision, would you vote for or against banning late-term abortion?"

[6]This item was one of the "roll-call" items. Respondents were first given a brief description of the roll-call vote on stem cell research, and were then asked "What do you think? If you were faced with this decision, would you vote for or against federal funds for this research?"

[7]This item reads: "President Bush recently spoke out in favor of a Constitutional Amendment defining marriage as strictly between a man and a woman. Do you support or oppose a Constitutional amendment banning gay marriage?" Respondents were then given a four-point Likert scale running from "strongly support" to "strongly oppose." This item was only asked of about half of the sample.

The analysis of the social items proceeds exactly as in the economic issues case (here, the identifying restrictions are that $\mu_{\text{Ohio}} = 0$ and the discrimination parameter on the abortion item is fixed to 1). Figures 3 and 4 give the results for the state mean and standard deviation parameters along this dimension.

The results in figures 3 and 4 largely parallel the results for the economic dimension. The reader should particularly note that much like on the economic dimension, there is a real limit to our ability to differentiate more or less heterogeneous states. But as we stressed in the paper, this does not render our results useless, it merely indicates the limits of our ability to draw confident inferences from the data.

Readers may be particularly curious to see how our estimates of ideological heterogeneity compare across dimensions. Figure 5 plots the estimates of state-level ideological heterogeneity on the social dimension against their counterparts from the social issues dimension.

As figure 5 reveals, there is some relationship between the estimates of heterogeneity across dimensions (the correlation is 0.36), suggesting that some heterogeneity is determined by structural factors common to both dimensions.[8] However, the figure makes plain that there is also clearly a dimension-specific component as well—there are some states that are highly heterogeneous on one dimension but not the other (e.g., NH, VT, HI, etc.). We leave exploring these differences for future work, simply noting here that there is some relationship between the two dimensions.

4 Description of the Demographic Variables

This section details the construction of each demographic measure. All of our demographic data is based on the 2000 U.S. Census. Some variables come directly from the Census

[8] Note that it is not meaningful to compare the scales of the two dimensions: they have different identifying restrictions, hence the different scales. So one cannot imply that states are overall more heterogeneous on the social dimension simply because the values of the standard deviation parameters are larger.

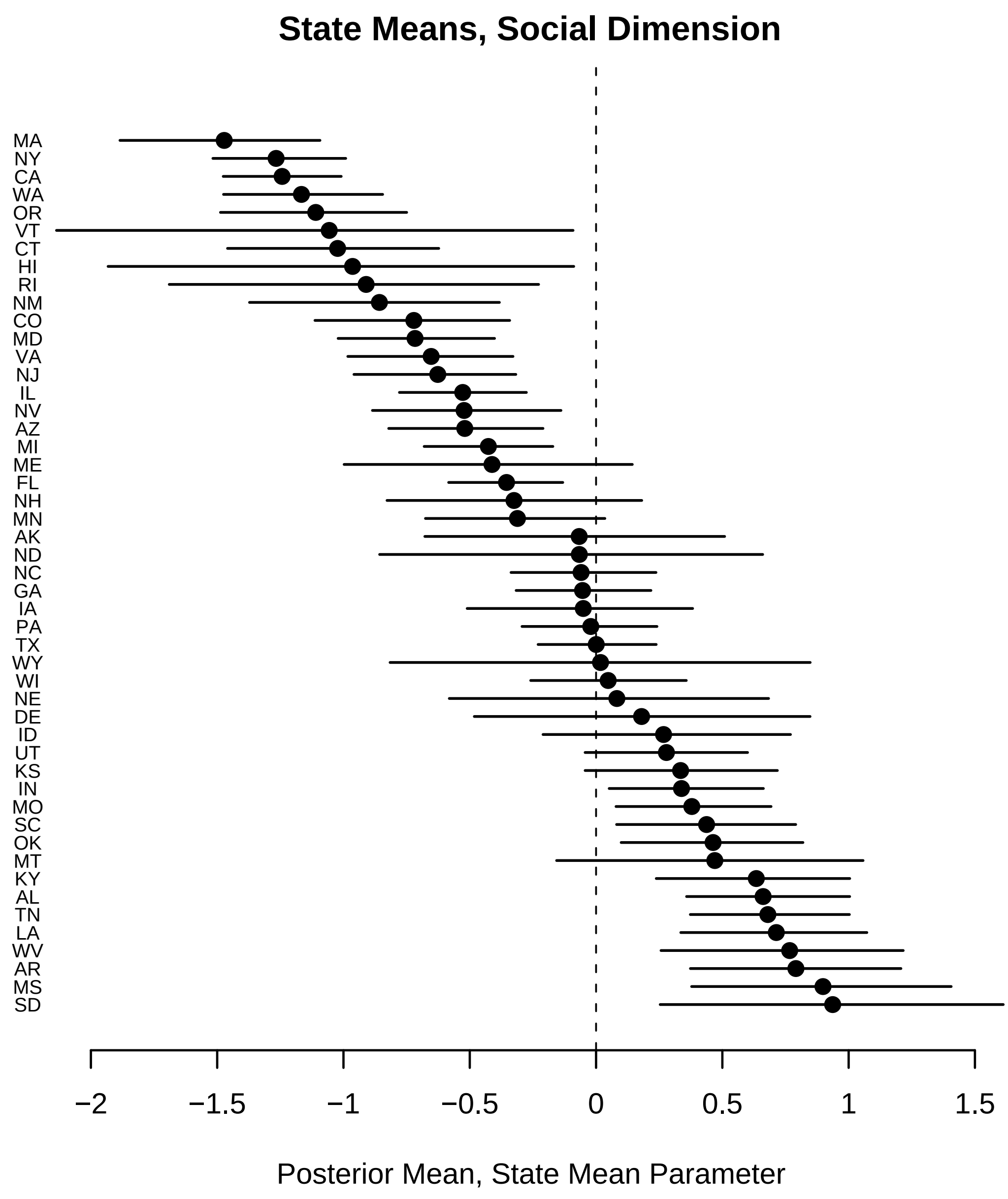

Figure 3: Posterior means and 95% posterior credible intervals for the state mean parameters, social issues dimension.

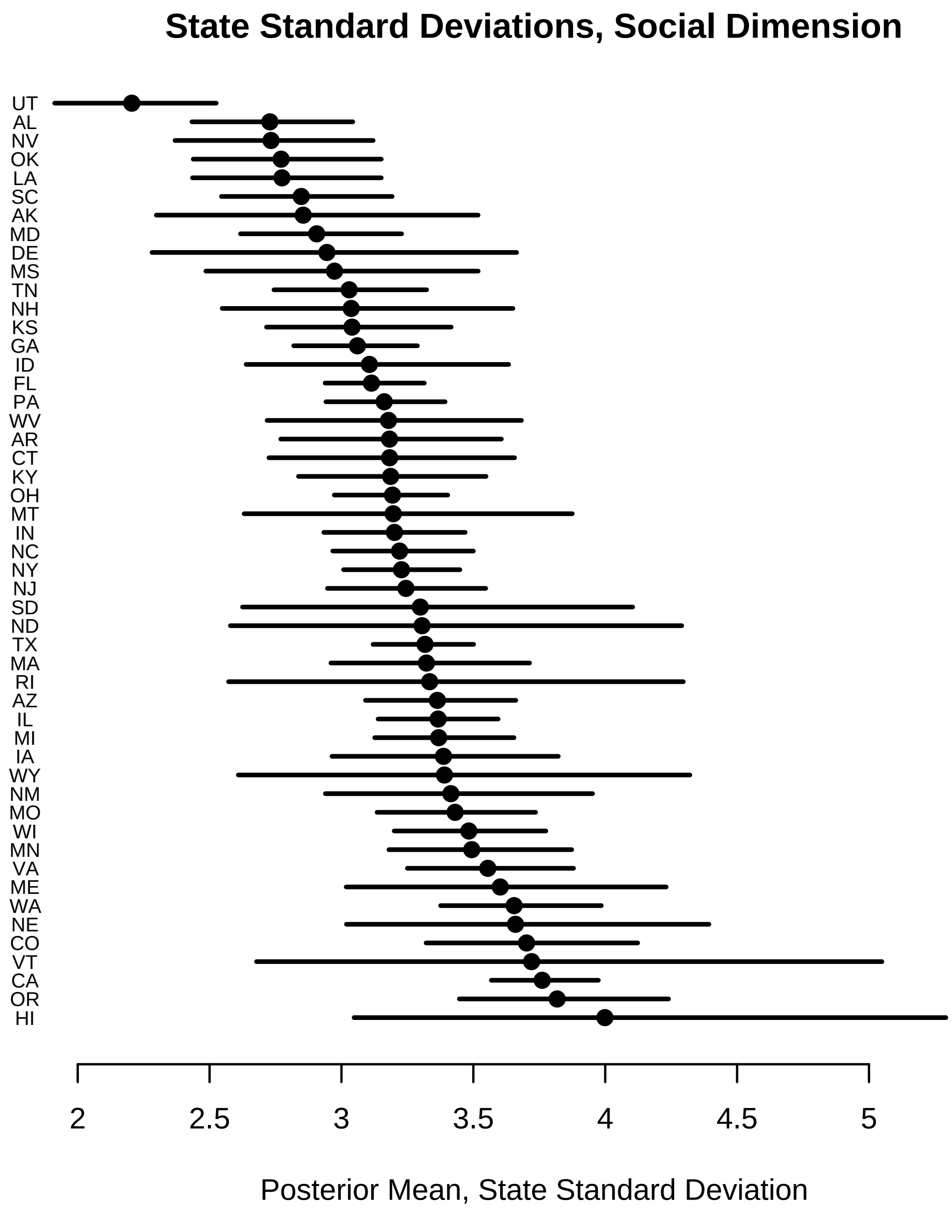

Figure 4: Posterior means and 95% posterior credible intervals for the state standard deviation parameters, social issues dimension.

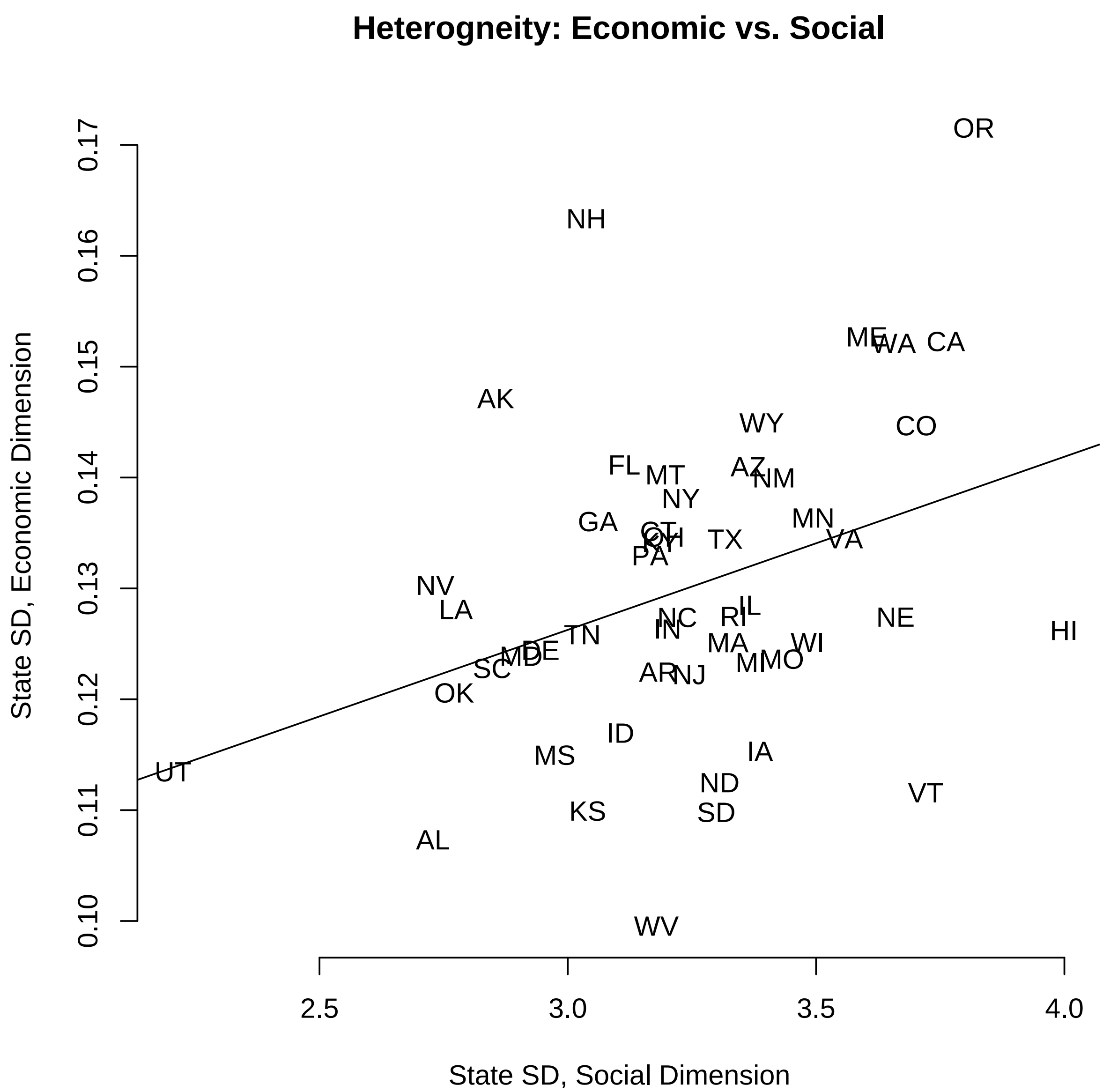

Figure 5: Comparison of the estimates of ideological heterogeneity on the social and economic dimensions.

itself (tables DP-1 through DP-4, available online at `http://censtats.census.gov`), while others come from the 2007 edition of the Statistical Abstract of the United States (available online at `http://www.census.gov/compendia/statab/`). For each variable below, we give the definition of the variable and the source.

- Percent African-American: Percent of the population identifying their race as "African-American" or "black." Source: 2000 US Census.[9]
- South: 1 if the state was a member of the former Confederacy, 0 otherwise.
- Percent Foreign Born: Percent of the population that is foreign born. Source: 2000 U.S. Census.
- Percent Urban: Percent of the population that lives in an urban area. Source: Statistical Abstract, table 33.
- Population Density: Population per square mile. Source: Statistical Abstract, table 18.
- Percent with a College Degree: Percent of the population (over age 25) that has at least a Bachelor's degree. Source: 2000 U.S. Census.
- Percent Homeowners: Percent of the occupied housing units that are owner-occupied. Source: 2000 U.S. Census.
- Percent Union Members: Percent of the population that belongs to a labor union. Source: Statistical Abstract, table 647.
- Percent White Collar: Percent of the population that is in "management, profession and related occupations" under the Census job classification scheme. Source: 2000 U.S. Census.
- Median Income: Median household income in the state. Source: 2000 U.S. Census.
- Total Population: Total state population. Source: 2000 U.S. Census.

Constructing the Sullivan Index

This is a brief explanation of how we re-calculated the Sullivan index for the 2000 Census (and associated) data. Sullivan's original 1973 measure had 6 components: education, income, occupation, housing status, ethnicity and religion. We added two additional variables

[9]Here, for simplicity, we focused on the overwhelming fraction of respondents who select only one race for themselves. This only excludes, at most, several percent of the respondents in any given state (typically only 1-2 percent).

for race and urbanicity (both of which were also used in Koetzle's 1998 political diversity index), and changed the categories of his measure to make it applicable to the modern era. All of the variables come from the 2000 Census and are described above, with the exception of the religion variables. The data on the number of Christians and Jews in each state come from the Statistical Abstract (table 75), and the data on Catholics comes from the Glenmary Research Center (see `http://www.glenmary.org/GRC/new/Catholic_data_states/50_states.htm` for more information). Our categories were:

1. Education:
   (a) Less than a high school diploma
   (b) High School Diploma or equivalent, but no further schooling
   (c) Some college (includes Associate's Degree)
   (d) Bachelors Degree or More
2. Income (using household income, 1999 dollars)
   (a) Less than 34,999 dollars per year
   (b) Between 35,000-99,999 per year
   (c) At least 100,000 per year
3. Occupational Variables
   (a) White-collar occupation
   (b) Other
4. Housing
   (a) Home owner
   (b) Renter
5. Ethnicity
   (a) Native born
   (b) Foreign born
6. Religion
   (a) Jewish
   (b) Catholic
   (c) Christian, not Catholic

(d) Other

7. Race

   (a) White, non-Latino/Hispanic
   (b) Black/African-American
   (c) Latino or Hispanic
   (d) Other Race

8. Urbanicity

   (a) Urban
   (b) Other

While some might argue that we should add further categories (e.g., adding a blue-collar measure to the occupation variable), we wanted to stick as closely as possible to the original measure. Our current configuration allows us to balance closeness to Sullivan's original measure with applicability to the modern era. Using these variables, we followed Sullivan's 1973 formula, given in his original text on pages 70-73.

Table of Correlations with Demographic Items

Table 2 gives the correlation between our measure of ideological heterogeneity on the economic dimension and various demographic measures used in the paper.

These correlations seem consistent with earlier findings in the literature (Lewis and Gerber 2004; Kuklinski 1977; Erikson, Wright, and McIver 1994) which make clear that demographics are *not* good proxies for opinion diversity at the state level (indeed, our measure of average state ideology is also only weakly correlated with demographics as well). Our findings extend this same conclusion to ideological heterogeneity. This is not to say that there is no relationship between demographics and ideological diversity (or average ideology), but our results suggest that demographics do not cleanly substitute for either the mean or variance of state ideology.

Variable	Standard Deviation	Party Distance
Log Proportion African-American	-0.14	-0.14
Log Median Income	0.32	0.31
Log Proportion Foreign Born	0.31	0.20
Log Proportion Urban	0.28	0.30
Log Population Density	-0.04	-0.05
Log Total Population	0..15	0.20
Log Proportion Bachelor's Degree	0.36	0.33
Log Proportion Homeowners	-0.33	-0.10
Log Union Membership	0.22	0.11
Log Proportion White Collar	0.27	0.27
Sullivan Index	0.27	0.21

Table 2: Correlation between various demographic variables and the measure of ideological heterogeneity.

5 Comparative Estimates of Heterogeneity

Table 3 below gives several different estimates of heterogeneity by state: the standard deviation and party distance measures developed in the body of the paper (using the CCES data), along with the variance of the CCES items used, and the Sullivan index of demographic heterogeneity. All of these measures (along with several other estimates of ideological heterogeneity) are given in downloadable files available at the authors' websites.

Table 3: Table giving various estimates of state ideological heterogeneity, see the text for more details.

State Abbreviation	Standard Deviation Measure	Party Distance Measure	Variance, Raw Items	Sullivan Index
AK	0.147	0.289	0.289	0.58
AL	0.107	0.193	0.258	0.553
AR	0.122	0.254	0.269	0.538
AZ	0.141	0.301	0.298	0.619
CA	0.152	0.313	0.298	0.67

continued on next page

Table 3: *continued*

State Abbreviation	Standard Deviation Measure	Party Distance Measure	Variance, Raw Items	Sullivan Index
CO	0.145	0.324	0.298	0.59
CT	0.135	0.271	0.28	0.59
DE	0.124	0.255	0.275	0.576
FL	0.141	0.289	0.289	0.632
GA	0.136	0.252	0.286	0.599
HI	0.126	0.183	0.254	0.628
IA	0.115	0.321	0.292	0.534
ID	0.117	0.267	0.271	0.547
IL	0.129	0.244	0.275	0.628
IN	0.126	0.241	0.279	0.553
KS	0.11	0.206	0.263	0.567
KY	0.134	0.233	0.273	0.532
LA	0.128	0.234	0.278	0.598
MA	0.125	0.292	0.271	0.6
MD	0.124	0.236	0.275	0.612
ME	0.153	0.277	0.287	0.486
MI	0.123	0.249	0.275	0.58
MN	0.136	0.345	0.299	0.558
MO	0.124	0.259	0.281	0.562
MS	0.115	0.208	0.27	0.551
MT	0.14	0.333	0.289	0.525
NC	0.127	0.248	0.278	0.563

continued on next page

Table 3: *continued*

State Abbreviation	Standard Deviation Measure	Party Distance Measure	Variance, Raw Items	Sullivan Index
ND	0.113	0.234	0.261	0.532
NE	0.127	0.206	0.277	0.563
NH	0.163	0.377	0.316	0.518
NJ	0.122	0.267	0.279	0.61
NM	0.14	0.308	0.292	0.623
NV	0.13	0.276	0.287	0.629
NY	0.138	0.258	0.275	0.657
OH	0.135	0.26	0.286	0.566
OK	0.121	0.222	0.275	0.568
OR	0.172	0.342	0.302	0.568
PA	0.133	0.234	0.283	0.573
RI	0.128	0.234	0.264	0.601
SC	0.123	0.279	0.285	0.567
SD	0.11	0.16	0.253	0.535
TN	0.126	0.256	0.282	0.556
TX	0.134	0.296	0.292	0.644
UT	0.113	0.221	0.263	0.553
VA	0.135	0.268	0.283	0.591
VT	0.112	0.157	0.242	0.487
WA	0.152	0.36	0.299	0.582
WI	0.125	0.274	0.287	0.562
WV	0.1	0.182	0.252	0.484

continued on next page

Table 3: *continued*

State Abbreviation	Standard Deviation Measure	Party Distance Measure	Variance, Raw Items	Sullivan Index
WY	0.145	0.282	0.281	0.548

6 Estimation and Inference via Markov Chain Monte Carlo

This section provides additional details that explain how we estimated our model. As we discussed in the body of the paper, we adopt a Bayesian approach for estimation and inference. Here, while it might be theoretically possible to estimate our model with frequentist tools (e.g., marginal maximum likelihood), doing so will be taxing to say the least (given the large number of parameters in our model). A Bayesian approach will likely prove to be simpler, although the final answers are unlikely to be driven by our choice of method.

Absent additional restrictions, the model as presented in the body of the paper is not identified owing to the fact that ideology has no natural metric (e.g., what does it mean to be one unit more liberal than someone else?). Rivers (2003) verifies that two restrictions are needed to identify a unidimensional model—one to set the location (e.g., the zero point) and one to set the scale. We do this by fixing the intercept and slope parameter for one of the items to be 0 and 1 respectively.[10] With these restrictions in place, the latent trait is identified and we can proceed with our analysis.

Before moving to a more detailed discussion of the model, we need to specify the prior

[10]Readers should not attach any substantive significance to these choices, they are simply identifying restrictions: the issue is akin to measuring temperature using the Fahrenheit or the Celsius scale (Jackman 2004b). Different restrictions yield different scales, but the underlying quantity of interest is the same.

distributions for the model parameters. For the item discrimination and difficulty parameters (i.e., β,α, λ, and ζ), we use a vague normal distribution, $\beta_j \sim N(0, 10)$,$\zeta \sim N(0, 10)$, $\alpha_j \sim N(0, 10)$, $\lambda_j \sim N(0, 10)$ $\forall j$ (where j indexes survey items).[11] For the threshold parameters, recall that $\tau_{j,k} = \sum_{l=1}^{K} \delta_{j,l}$ (j indexes survey items, k indexes response options), so we place a prior on τ by placing one on δ. Here, $\delta_{j,1} \sim N(0, 100)$, and $\delta_{l,m} \sim exp(2)$ $\forall$ $l \geq 2$. Finally, recall that $x_i \sim N(\mu_{m[i],p[i]}, \sigma^2_{m[i],p[i]})$, where i indexes individuals, m indexes states, and p indexes political parties. Hence, $m[i]$ gives the state in which respondent i is a resident and $p[i]$ gives respondent i's party affiliation. Here, $\mu_{m,p} \sim N(0, 1)$ and $\sigma_{m,p} \sim U(0, 25)$.

In a Bayesian analysis, we want to compute the posterior density of all model parameters.[12] Then let $\boldsymbol{\omega} = \{\boldsymbol{x}, \boldsymbol{\alpha}, \boldsymbol{\tau}, \boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\zeta}, \boldsymbol{\lambda}\}$ represent the parameters of interest in our model, and let $\mathbf{Y}$ be the data matrix of individual's policy positions, made by stacking the relevant data across individuals and items. We want to estimate $p(\boldsymbol{\omega}|\mathbf{Y})$, that is, the joint posterior density of all model parameters conditional on the data. Via Bayes' Rule, this posterior is proportional to the prior density times the likelihood: $p(\boldsymbol{\omega}|\mathbf{Y}) \propto \mathcal{L}(\mathbf{Y}|\boldsymbol{x}, \boldsymbol{\alpha}, \boldsymbol{\tau}, \boldsymbol{\beta}, \boldsymbol{\zeta}, \boldsymbol{\lambda})P(\boldsymbol{\omega})$, where $\mathcal{L}(\mathbf{Y})$ is the likelihood for $\mathbf{Y}$, and $p(\boldsymbol{\omega})$ is the prior density over the model parameters.

The MCMC algorithm generates a random tour of the model posterior density. MCMC algorithms make sampling from the high-dimensional joint posterior simpler by breaking down the joint posterior density into the conditional densities and sampling from them in turn (see Jackman (2000, 2004a) for a review of the MCMC methods and applications in political science; see also Gelman et al. (1995) on Bayesian methods more generally).

For our model, then, iteration t of the MCMC algorithm involves sampling from the

[11]Technically, α, β & ζ, and λ correspond to items from the ordinal, binary, and continuous sections (respectively) of our model. As such, they should have separate indices, but in the interest of simplicity, we use a common index (j) throughout. We also suppress the ω^2 term (the variance for the continuous items) in the interest of simplicity.

[12]We assume we have no missing data in our discussion of the model posterior. Again, we could carry around notation for missing data, but it would add additional notation without much gain in clarity.

following sets of conditional distributions:

1. sample $x_i^{(t)}$ from $p_x(x_i|\boldsymbol{\alpha}^{(t-1)}, \boldsymbol{\lambda}^{(t-1)}, \boldsymbol{\beta}^{(t-1)}, \boldsymbol{\zeta}^{(t-1)}, \boldsymbol{\tau}^{(t-1)}, \boldsymbol{\mu}^{(t-1)}, \boldsymbol{\sigma}^{(t-1)}, \mathbf{Y})$, $i = 1, \ldots, n$

2. sample $\alpha_j^{(t)}$ from $p_\alpha(\alpha_j|\boldsymbol{x}^{(t)}, \boldsymbol{\lambda}^{(t-1)}, \boldsymbol{\beta}^{(t-1)}, \boldsymbol{\zeta}^{(t-1)}, \boldsymbol{\tau}^{(t-1)}, \boldsymbol{\mu}^{(t-1)}, \boldsymbol{\sigma}^{(t-1)}, \mathbf{Y})$, $j = 1, \ldots, J$

3. sample $\lambda_j^{(t)}$ from $p_\lambda(\lambda_j|\boldsymbol{x}^{(t)}, \boldsymbol{\alpha}^{(t)}, \boldsymbol{\beta}^{(t-1)}, \boldsymbol{\zeta}^{(t-1)}, \boldsymbol{\tau}^{(t-1)}, \boldsymbol{\mu}^{(t-1)}, \boldsymbol{\sigma}^{(t-1)}, \mathbf{Y})$, $j = 1, \ldots, J$

4. sample $\beta_j^{(t)}$ from $p_\beta(\beta_j|\boldsymbol{x}^{(t)}, \boldsymbol{\alpha}^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{\zeta}^{(t-1)}, \boldsymbol{\tau}^{(t-1)}, \boldsymbol{\mu}^{(t-1)}, \boldsymbol{\sigma}^{(t-1)}, \mathbf{Y})$, $j = 1, \ldots, J$

5. sample $\zeta_j^{(t)}$ from $p_\zeta(\zeta_j|\boldsymbol{x}^{(t)}, \boldsymbol{\alpha}^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{\beta}^{(t)}, \boldsymbol{\tau}^{(t-1)}, \boldsymbol{\mu}^{(t-1)}, \boldsymbol{\sigma}^{(t-1)}, \mathbf{Y})$, $j = 1, \ldots, J$

6. sample $\boldsymbol{\tau}_j^{(t)}$ from $p_\tau(\boldsymbol{\tau}_j|\boldsymbol{x}^{(t)}, \boldsymbol{\alpha}^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{\beta}^{(t)}, \boldsymbol{\zeta}^{(t)}, \boldsymbol{\mu}^{(t-1)}, \boldsymbol{\sigma}^{(t-1)}, \mathbf{Y})$, $j = 1, \ldots, J$

7. sample $\mu_{m,p}^{(t)}$ from $p_\mu(\mu_{m,p}|\boldsymbol{x}^{(t)}, \boldsymbol{\alpha}^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{\beta}^{(t)}, \boldsymbol{\zeta}^{(t)}, \boldsymbol{\tau}^{(t)}, \boldsymbol{\sigma}^{(t-1)}, \mathbf{Y})$, $m = 1, \ldots, M$, $p = 1, 2, 3$

8. sample $\sigma_{m,p}^{(t)}$ from $p_\sigma(\sigma_{m,p}|\boldsymbol{x}^{(t)}, \boldsymbol{\alpha}^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{\beta}^{(t)}, \boldsymbol{\zeta}^{(t)}, \boldsymbol{\tau}^{(t)}, \boldsymbol{\mu}^{(t)}, \mathbf{Y})$, $m = 1, \ldots, M$, $p = 1, 2, 3$

Given the model and prior densities described above, there is no conjugacy stemming from the fact that we have ordinal items and normal parameter distributions. As a consequence, the conditional distributions needed to implement the MCMC algorithm are non-standard, and sampling from them will be more difficult. Using winBUGS (Spiegelhalter et al. 2003), however, makes analyzing this problem much simpler because of its use of sophisticated algorithms such as slice sampling (Neal 2003).

Here, we initialized the MCMC algorithm with zeros for all parameters except the δ parameters, which are started at 2, and the latent trait parameters, which were initialized at the individual's value on the liberal-conservative self-identification scale, rescaled so that 0 was the "moderate" position. we ran the algorithm for 66,000 iterations, and discarded the first 6,000 iterations as a burn-in period to ensure that the algorithm had moved away from the initial values. After discarding the burn-in observations, we save every 30th iteration, resulting in 2000 approximately independent draws from the joint posterior density that are then used for the analysis reported in the text.

We make three additional choices worth noting. First, for the sake of simplicity, we treat missing data as missing data and do not perform any imputations, though one could use the MCMC algorithm for that purpose (for an example, see Levendusky, Pope, and Jackman 2008). Second, we also include all respondents to the common content module of the CCES, rather than restricting the sample to voters only. Finally, we analyze the unweighted CCES sample. We do so because the weights provided with the CCES are (poststratification) weights to make the sample nationally representative (e.g., to come as close as possible to the demographic profile of the ACS), not representative at the state level (or, more relevant for our purposes, partisan sub-groups within each state). Given this, we choose to analyze the unweighted data, and leave the question of state-level weights for future research.

7 2000 Annenberg National Election Study

The paper also uses results from the 2000 National Annenberg Election Study (Romer et al. 2004). Specifically, we used the national rolling cross-section data from 18 July 2000 until 6 November 2000. Why did we limit the data in this way? Although the Annenberg rolling cross-section study covers a longer period (from 14 December 1999 until 19 January 2001), using data from before July or after the election is problematic. Before July, most respondents were probably not particularly attuned to the campaign, and as such, their attitudes would likely differ from respondents interviewed later during the height of the campaign (where people would be learning where their party stood on the issue and adjusting their own attitudes accordingly). Using data from after the campaign could also be problematic, since the disputed election might color their attitudes. We felt that selecting the July to November sample would allow us to have enough data to actually estimate the model with still retaining a relatively homogeneous sample of voters.

There is a limitation to the Anneberg data, however. The rolling cross-section data has the nation as its sampling frame and not the individual states, so the study may not reproduce representative samples at the state level. However, a number of previous scholars have used these data to measure state-level preferences, so this does not seem to be a particularly troubling assumption (Clinton 2006; Johnston, Hagen, and Jamieson 2004). Given this, we assume that the Annenberg samples are (at least approximately) representative of the state-level electorates.

Much like the CCES data, we selected items that measured respondent's preferences about the size of government and its role in the economy. Specifically, we used the following items: (1) whether or not the respondent felt taxes should be cut, (2) whether or not the inheritance tax should be eliminated, (3) whether or not a flat tax should be introduced, (4) whether or not funds from social security should be invested in the stock market, (5) whether or not military spending should be increased, and (6) whether or not spending on the poor should be increased. Dimensional analyses reveal that these items form a one-dimensional structure.

The model and analysis proceeds as described above for the CCES data. Here, we replicate table 1 from the paper, which gives the relationship between demographics and the measures of heterogeneity. Table 4 gives the results.

The results in table 4 reinforce those from the body of the paper: demographics do not do a good job (either individually or collectively) predicting ideological heterogeneity.

8 Comparing the NAES and CCES Measures of Heterogeneity

In the paper, we report two sets of estimates of ideological heterogeneity: one from the CCES, and another from the NAES. Figure 6 compares them to one another (for both the standard deviation and the between-party distance metrics).

Variable	Standard Deviation	Party Distance
Intercept	8.18	-17.59
	(6.37)	(12.46)
Log Proportion African-American	0.02	-0.01
	(0.06)	(0.12)
South	0.65	-0.37
	(0.62)	(1.21)
Log Median Income	-0.29	0.91
	(0.57)	(1.11)
Log Proportion Foreign Born	0.09	0.01
	(0.07)	(0.13)
Log Proportion Urban	-0.10	-0.17
	(0.37)	(0.73)
Log Population Density	0.05	-0.11
	(0.05)	(0.09)
Log Total Population	**-0.10**	**0.24**
	(0.05)	(0.10)
Log Proportion Bachelor's Degree	0.28	1.05
	(0.85)	(1.64)
Log Proportion Homeowner	-0.49	1.22
	(0.73)	(1.43)
Log Proportion Union Membership	-0.04	0.11
	(0.10)	(0.19)
Log Proportion White Collar	-0.38	-0.66
	(1.16)	(2.26)
Sullivan Index	0.13	1.28
	(2.59)	(5.07)
Log Proportion African-American × South	-0.21	0.20
	(0.20)	(0.40)
N	48	48
R^2	0.21	0.41
F-statistic	0.72	2.09
p-value	0.73	0.05

Table 4: Replication of table 1 from the paper (relationship between heterogeneity measures and demographics) using measures from the 2000 NAES data.

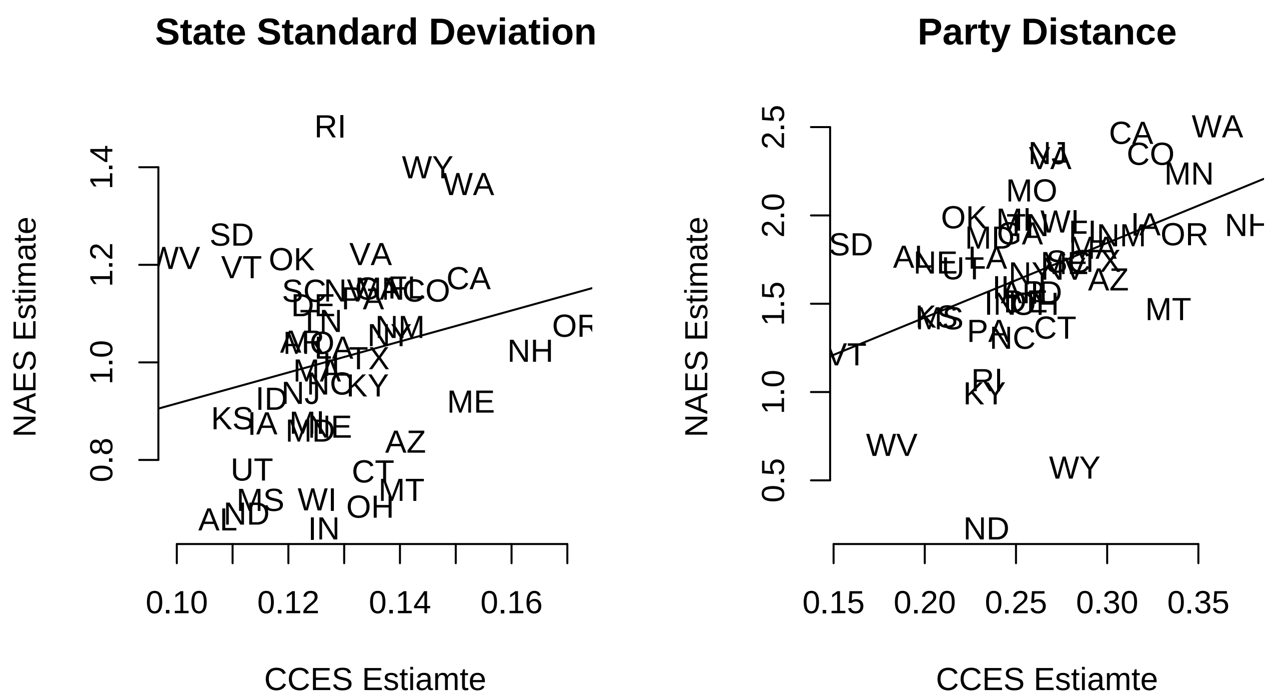

Figure 6: Heterogeneity estimates produced from the CCES and NAES data for the fifty states.

Figure 6 demonstrates that the two measures are moderately related: the standard deviation measures are correlated at 0.23, the party distance measure at 0.45. However, one would not expect these two estimates to be very highly correlated. While they both tap the same underlying theoretical construct (state ideological heterogeneity), there are a host of important differences between them:

1. Year of the survey: the NAES data comes from 2000, the CCES comes from 2006. The surveys were done in different political climates, which will undoubtedly influence the results.

2. Sampling Frames: the CCES uses an opt-in Internet sample (adjusted on the back-end to make it demographically representative), the NAES rolling cross section is a RDD telephone sample.

3. Items used: while the items used are similar and tap the same latent dimension (by construction), they are not identical, which will change the results.

In the end, there are just too many differences between the studies to expect them to be perfectly correlated. One can make a parallel to ideal point estimates for legislators. Imagine that we took random samples of 10 votes in different years and generated ideal point estimates using only those 10 votes. We would not be surprised to see that the ideal point estimates differed (perhaps by a lot) across these different samples. The same reasoning applies here. We should expect the different sample estimates to be correlated, but they should *not* be identical.

References

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

Erikson, Robert, Gerald Wright, and John McIver. 1994. *Statehouse Democracy: Public Opinion and Policy in the American States*. New York: Cambridge University Press.

Gelman, Andrew, John Carlin, Hal Stern, and Donald Rubin. 1995. *Bayesian Data Analysis*. London, UK: Chapman and Hall.

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

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

Jackman, Simon. 2004b. "What Do we Learn from Graduate Admissions Committees? A Multiple Rater, Latent Variable Model, with Incomplete Discrete and Continuous Indicators." *Political Analysis* 12: 400-424.

Johnston, Richard, Michael Hagen, and Kathleen Jall Jamieson. 2004. *The 2000 Presidential Election and the Foundations of Party Politics*. New York: Cambridge University Press.

Koetzle, William. 1998. "The Impact of Constituency Diversity upon the Competitiveness of U.S. House Elections, 1962-96." *Legislative Studies Quarterly* 23: 561-73.

Kuklinski, James. 1977. "Constituency Opinion: A Test of the Surrogate Model." *Public Opinion Quarterly* 41: 34-40.

Levendusky, Matthew, Jeremy Pope, and Simon Jackman. 2008. "Measuring District-Level Partisanship with Implications for the Analysis of U.S. Eelctions." *Journal of Politics* 70 (3): 736-753.

Lewis, Jeffrey, and Elisabeth Gerber. 2004. “Beyond the Median: Voter Preferncces, District Heterogeneity, and Representation.” *Journal of Political Economy* 106: 1364-1383.

Neal, Radford. 2003. “Slice Sampling (with Discussion).” *Annals of Statistics* 31: 705-767.

Rivers, Douglas. 2003. “Identification of Multidimensional Spatial Voting Models.” Manuscript: Stanford University.

Romer, Daniel, Kate Kenski, Paul Waldman, Christopher Adasiewicz, and Kathleen Hall Jamieson. 2004. *Capturing Campaign Dynamics: The National Annenberg Election Study.* New York: Oxford University Presss.

Spiegelhalter, David, Andrew Thomas, Nicky Best, and Andrew Lunn. 2003. “WinBUGS User Manual Version 1.4.” Cambridge, UK: MRC Biostatistics Unit.

Sullivan, John. 1973. “Political Correlates of Social, Economic and Religious Diversity in the American States.” *Journal of Politics* 35: 70-84.