Latent Versus Self-Reported Ideology

Benjamin E. Lauderdale
blauderd@princeton.edu, 609-258-0091

Princeton University
Department of Politics, Corwin Hall
Princeton, NJ 08544-1013

July 6, 2007*

ABSTRACT

Spatial voting axes are seldom directly measurable. In general, political scientists employ latent variable models that aim to infer the presence and characteristics of axes from their impact on measurable quantities such as voting behavior. Ideology in the electorate is a rare case where researchers regularly attempt to measure a spatial quantity directly, typically by asking people to rank themselves on ordinal scales from liberal to conservative. These self-reported ideology scores have been widely used as a nationally comparable measure of ideological position, but whether different citizens are using consistent criteria for assessing and reporting their ideology is largely unknown. To better understand the relationship between issue positions and ideology in the electorate, I develop a item-response model for self-reported ideology as a function of issue positions. I find that political information is required for voters to self-report ideology consistently with their issue positions. Groups which are less attuned to the terminology of the national political discourse tend to self-report ideology in ways that are less informative about their issue positions on the primary political axis defined by that discourse.

*The author thanks Larry Bartels and Chris Achen for reading early drafts as well as John Londregan, Kosuke Imai, and Jeff Gill for technical advice.

1. INTRODUCTION

Empirical spatial models of voting choice have been highly successful at explaining voting behavior in the U.S. Congress (Poole & Rosenthal 1997, Clinton, Jackman & Rivers 2004, Poole 2005) and Supreme Court (Martin & Quinn 2002). As political professionals, congressmen have strong personal and electoral incentives to develop coherent policy agendas and to behave consistently while judges similarly have consistent ideological commitments conditioning their rulings. Studies of Congress have observed that a strong first dimension is the dominant predictor of voting behavior in recent years. One can predict 90% of roll call votes in recent Congresses with a one dimensional model (Poole & Rosenthal 1997, Clinton, Jackman & Rivers 2004).

Not so the opinions of the American people. The relative instability and difficulty of measuring substantive policy positions in the American electorate has been repeatedly documented by political scientists for the last fifty years (Converse 1964, Achen 1975, Zaller 1992). In part because of the inconsistency of measured policy positions in the electorate, far less work has been done modeling voters' policy preferences spatially than has been done for political professionals. A notable exception is the work of Enelow & Hinich (1984), which models vote choice in terms of a policy axis and a valence axis using feeling thermometers to measure positions. The authors of *The American Voter* extensively explore the relationship of expressed issue positions to partisanship, but while their analysis of the structure of issue attitudes uses the liberal-conservative construct, it focuses on various scaling techniques rather than self-reports (Campbell, Converse, Miller & Stokes 1960, 188-215). Converse & Markus (1979) find that self-reported ideology is more stable over time for individuals than individual policy attitudes, but that this is in part due to the sizable fractions of citizens who claim no opinion and refuse to answer the question and those who steadfastly report themselves to be moderates. Knight (1985) shows that only among the most political informed citizens is self-reported ideology a predictor of candidate evaluation and voting behavior.

Despite these difficulties, the need for high quality measures of citizens' positions is acute. Most empirical analyses of the responsiveness of democratic officials to their constituents rely on the measurement of constituent preferences (Achen 1978).

In addition to the substantive constraints on spatial modeling of the electorate imposed by citizen incoherence, technical constraints have also limited research efforts. First, as the explanatory power of the spatial model declines, classical maximization techniques increasingly have convergence problems, a problem solved by transitioning to Bayesian techniques more appropriate to problems where the data are weakly informative (Jackman 2001). Second, spatial modeling requires large numbers of questions and respondents for statistical consistency (Haberman 1977, Londregan 2000) and the former are typically lacking in available data, though random-effects models can be used in some circumstances to leverage large numbers of respondents to compensate for small numbers of questions (Lewis 2001). Third, methods for extending the binary choice spatial models used for roll call voting to polychotomous choice models useful for ordinal polling questions have only recently been developed (Treier & Jackman 2002, Quinn 2004).

Because of spatial modeling's steep demands for data, few surveys are sufficiently broad to allow this sort of analysis; however, political scientists often need to have a estimate of an individual's position to control for population heterogeneity or to estimate variation in citizen preference across groups. Self-reported ideology is used widely in these circumstances, typically in the form of questions which asks citizens to rate their position from liberal to conservative on an ordinal scale with 3, 5 or 7 positions. Political scientists have long been concerned about the possibility of scale incomparability (Brady 1985) and developing methodology to improve comparability is an active area of research (King, Murray, Salomon & Tandon 2004, King & Wand 2007). In many data sets though, and with scales like ideology that are difficult to define clearly enough to provide useful baselines to respondents, authors have little choice but to use these ordinal scales as given. The default assumption is often to treat these data as "directly comparable" (Clinton 2006) and so studies assessing the

responsiveness of elected officials to constituent opinion (Clinton 2006, Bartels 2007) have used aggregated self-reports by congressional district or state. If there are systematic biases in the way that citizens self-report their ideologies, or certain groups of citizens self-report in ways uncorrelated to any substantive beliefs, there is further ambiguity in which measure of responsiveness we ought to use to assess the normative success of democratic institutions (Achen 1978). Because there is error in both self-reports of ideology and in self-reports of issue positions, a failure of self-reported ideology to reflect self-reported issue positions need not imply that ideology is a useless construct for understanding how citizens organize their political beliefs. Authors have often conceptualized self-reported ideology not as a summary of issue positions, but as a world-view from which certain issue positions will spring if an individual is sufficiently aware to realize the implications of their more fundamental, ideological position. To the extent that ideology fails to match issue positions, the question of which is at error is intrinsically difficult to answer. The analysis in this paper cannot address this question directly. It aims instead to assess to what extent self-reported ideology does in fact reflect issue positions and draw what conclusions are possible.

Recent research has begun to reassess the relationships between issue positions and ideology in the electorate by employing contemporary advances in political methodology. Treier & Jackman (2002) introduce an ordinal item response theory model to assess the relative discriminatory power of different approaches to uncovering ideology on a seven point scale. Treier & Hillygus (2005) use the same ordinal item response theory (IRT) model to estimate positions in one and two dimensions for members of the 2000 National Election Studies (NES) cross-section and then explore the role of cross-cutting ideological commitments on presidential vote. The authors find evidence in support of the multidimensionality that most popular political commentary assumes: the existence of an economic axis reflecting beliefs about redistribution and government involvement in the economy and a social axis reflecting the extent to which government ought to encourage or deprecate particular social behaviors through legislation. The authors also find that self-reported ideology is predicted by a one

dimensional IRT model for ordinal response using data in the 2000 National Election Study (NES), however they do not test whether it does so consistently across the electorate.

I extend the Treier and Hillygus model by estimating a heteroskedastic IRT model where the variances of individual spatial voting errors are estimated rather than assumed to be identical. The heteroskedastic model is especially useful in this context because while legislators and other political professionals can be expected to vote in highly organized (spatial) ways, citizens will organize their views to wildly varying degrees. Additionally, the heteroskedastic model uses the fact that some respondents are providing less information about the location of the axis (because they are not responding spatially) to reduce the weights on their responses in estimating the question parameters, yielding a modest efficiency gain. The estimated errors themselves give additional information about citizen beliefs: the extent to which we can describe their issue positions as spatial based on the primary axis through the population's beliefs. I estimate this model on the last four presidential election year NES, demonstrating that the primary axis through American public opinion has hardly moved in that 12 years. Employing these estimates of issue positions for each survey respondent in that period, I then model the process by which individuals generate responses to the self-reported ideology question as a function of these issue positions and demographic characteristics. There is enormous variation in the extent to which self-reported ideology reflects issue positions across individuals and the primary source of this variation appears to be the degree to which an individual is attuned to the terminology of national politics.

2. A MODEL FOR IDEOLOGY SELF-REPORTS

It need not be the case that the dimension that provides the maximum explanatory power for expressed policy preferences is the same axis as the liberal-conservative axis that self-reported ideology questions request. If it were the case that the liberal-conservative dimension reflected some particular, generally known axis in the higher dimensional issue space, it might

not be the axis which explained the most variation in polling responses. But regardless of whether there could be general agreement on a coherent definition of the relevant labels, it is surely the case that perceptions in most of the electorate of what constitutes being a liberal and a conservative are determined by the orientations of actual individuals rather than theoretical constructs.[1] Consequently, when we ask members of the electorate for their liberal-conservative positions—to the extent that we are asking for information about their policy positions at all—we are asking them to position themselves with respect to their fellow citizens, which is what estimators of the spatial model do statistically.

Setting aside the question of whether self-reported ideology ought to reflect spatial position, assume that citizens want to use their issue positions as a component of their self-reported ideology. To report on a one-dimensional ordinal scale, an individual would need several pieces of information. First, they need to know their own position on the relevant ideological axis. Second, since self-reported ideology questions imply a relative position—there is no absolute meaning to being a moderate—an individual must know where the center of the ideology space is. Third, the individual must know the spacing between the categories offered as response options which tell them how far from center they must be before their moderate "4" turns into a slightly conservative "5" or a very liberal "1". With these three pieces of information, the voter can determine their distance from the center and how many categories from the central category this spatial distance corresponds to. It is worth noting that all three of these requirements are preceded by an even more challenging intellectual calculation: the location of the primary political axis in the higher dimensional issue space.

The most straightforward statistical specification for this problem has citizens choosing the category of their response using an ordered probit model to turn latent relative position r_i^* into a survey response r_i in one of the question's k_{max} categories. Where the $k_{max}-1$ cutpoints between the categories are given by θ_1, θ_2, ..., $\theta_{k_{max}-1}$ with $\theta_0 = -\infty$ and $\theta_{k_{max}} = \infty$, the

[1] E.g., self-labeled conservatives are more likely to express what were once called liberal economic positions—free markets and minimal economic regulation—indeed these positions are currently called "liberal" in Europe.

categorical response is given by:

$$r_i = k \iff \theta_{k-1} < r_i^* \leq \theta_k$$

I allow individuals to use the scale differently: modeling the relative ideological position r_i^* as a linear function of the estimated spatial position z_i. Thus, individual i combines their location on the ideological scale z_i with a parameter κ_i to determine the center of the scale and uses a scaling parameter λ_i to determine the sensitivity of the self-reported ideological scale to their spatial position z_i:

$$r_i^* = \kappa_i + \lambda_i z_i + \epsilon_i$$

Individuals with $\lambda_i = 0$ fail to report their ideology in ways that depend on their spatial position. Positive λ_i correspond to variably responsive reporting. I do not constrain the stochastic elements to be $N(0, 1)$, but instead leave the error variance to be estimated, $\epsilon_i \sim N(0, \omega^2)$.[2] In effect, non-spatial contributions to self-reports are subsumed into the error term and κ_i.

We cannot estimate κ_i and λ_i for every individual since we only observe one realization of r_i for each z_i. These parameters are instead modeled using variables capturing individual demographic and behavioral attributes. This is equivalent to fitting an ordered probit model for r_i as a function of z_i, demographics and their interactions, where κ_i is the sum of the intercept and all non-interactive effects, and λ_i is the sum of the coefficient on z_i and all interactions between other variables and z_i. This approach allows the model to capture variation both in where individuals believe the political center is and in the extent to which deviations in spatial position imply deviations in ideology.[3]

[2] Model identification is described below.

[3] While changes in λ_i describe multiplicative scalings and κ_i additive scalings to the latent scale, I assume that all individuals use the same relative cutting line θ_k spacing, though conceivably there might be demographic variation in that as well.

2.1. *Estimating Spatial Position from Ordinal Polling Response*

The above model for self-reported ideology depends on spatial position z_i, which itself must be estimated. I employ a heteroskedastic variant of the ordinal generalization of the two parameter probit IRT model (Treier & Jackman 2002, Quinn 2004). The realized ordinal response y_{ij} of individual i to issue polling question j is given in terms of the latent response y^*_{ij}:

$$y_{ij} = g \iff \gamma_{g-1} < y^*_{ij} \leq \gamma_g$$

The cutpoints γ_g are assigned an improper uniform prior. The latent response is given by the usual function of the question parameters α_j and β_j,

$$y^*_{ij} = \alpha_j + \beta_j z_i + \epsilon_{ij}$$

The priors on the z_i are $N(0, 1)$. Where this specification diverges from those previously used is that it allows different voters to make spatial voting errors of different magnitudes, $\epsilon_{ij} \sim N(0, \sigma_i)$.

Likelihood maximization techniques tend to fail when faced with highly parameterized models such as those described by IRT, so draws from the posterior distribution of the model is simulated by Markov Chain Monte Carlo (MCMC). Previous authors have demonstrated that IRT models are amenable to Gibbs sampling (Clinton, Jackman & Rivers 2004, Treier & Hillygus 2005), so that approach is adopted.

The performance of Gibbs sampling approaches to ordinal response models which draw each parameter singly can be very poor because strong posterior correlation of the cutpoint locations and the latent position leads to slow convergence.[4] This is a problem here because of the computational expense of the IRT model. With n respondents and m questions, on

[4]Perversely, this problem is exacerbated by having more respondents (Cowles 1996).

the order of $n \cdot m$ quantities must be drawn randomly from their conditional distribution for each iteration of the sampler, $\approx 2000 \cdot 15 = 30,000$ for single NES studies. A convenient solution to this problem is to jointly draw the cutpoint locations and the latent quantity from their joint posterior (Cowles 1996). Because their joint distribution does not have convenient conjugacy properties, this sampling must be done by a Metropolis-Hastings step within the Gibbs sampler. Cowles (1996) demonstrates that with this modification, convergence occurs even from poor starting values in less than 1000 iterations and mixing is rapid, allowing shorter samples thereafter.

Appendix A gives the sampling scheme for the posterior distribution of the model parameters. The IRT model proceeds in four stages, first jointly sampling the latent utilities y_{ij}^* and cutpoints γ_g by a Metropolis-Hastings step, then the question parameters α_j, β_j, the latent positions z_i, and finally the σ_i. Having drawn new values of the latent positions from their posterior distribution, the sampler can proceed to the model for self-reported ideology. As before, the latent ideologies r_i^* and cutpoints θ_k are jointly sampled using a Metropolis-Hastings step. The ordered probit coefficients κ, λ are sampled using a multivariate normal Gibbs step, and the variance of the latent variable error term is drawn from an inverse gamma distribution.

In this particular case, because the model for self-reported ideology is one-way—the posterior draws of the spatial positions z_i are used to estimate the ordered probit parameters, but not vice versa—it is possible to draw and store the z_i for use in the latter stages of the Gibbs sampler. Since estimating this model via compiled code is necessary to achieve tolerable performance, I have modified the C++ code written by Martin & Quinn (2007) to simulate the ordinal IRT model. Since generating the z_i chain is most of the computational burden for this problem—and is generally so for second stage analyses using ideal point estimates as independent variables—this approach yields the additional benefit that one can rapidly explore multiple model specifications for the second stage with a single stored z_i

chain.[5]

To identify the model requires fixing the latent scales in both the item-response model and in the ordered probit model. In the IRT portion of the model, I constrain the positions with a $N(0,1)$ prior.[6] To identify the σ_i I supply a diffuse conjugate prior and constrain $\sum_i^n \frac{1}{\sigma_i} = n$ via renormalization. In the ordered probit model for self-reported ideology, I center the scale by fixing the central two cutpoints θ_3 and θ_4 to the values $-1/2$ and $1/2$ and estimating ω^2. I constrain two cutpoints instead of constraining σ^2 plus one cutpoint for cross-model comparability and parameter interpretation. Setting the moderate ideology category to have a width of one and a center at zero allows us to transparently interpret the relative size of the categories, the magnitude of the coefficients relative to the category widths, and relative magnitudes of the deterministic and stochastic components of the model. Moreover, it guarantees that when we change model specifications, we can make direct comparisons of the coefficient magnitudes because the scale itself remains fixed.[7] Mathematical details of the Metropolis-Hastings-in-Gibbs sampler are provided in Appendix A. Simulations described in this paper used a burn-in of 4000 iterations followed by a posterior sample of 1000 iterations with no thinning.[8]

3. RESULTS AND DISCUSSION

To fit the model described above, I use data from the 1992-2004 NES studies, taken from the ANES cumulative file (2005). The questions used are listed in Table 1 and are all asked on the 1992, 1996, 2000, and 2004 surveys. Off-year and earlier elections are excluded because

[5]Of course, having settled on a second-stage specification, the analyst must ensure that results do not change with a different sample from the posterior of z_i.

[6]This leaves the model unidentified as to the sign of z_i. I set the signs of z_i and the β_j such that conservative positions are positive to match the polarity of the self-report scale and convention.

[7]While the outer categories could potentially shift, this would require rather odd patterns in the data. Such patterns do not seem to exist: there is little change in the outer cutpoints across specifications.

[8]Longer simulations on individual NES yield posterior sample summaries that are indistinguishable from those described. Longer simulations on the pooled NES sample were impossible with the code employed because of address space limitations.

the list of consistently asked questions becomes too short and unrepresentative of the sphere of live political issues. I focus on these four surveys to maximize comparability. I fit the model separately for each survey to track changes in the relationship between individual policy areas and the primary axis over time. This provides a convenient check on my results regarding group differences in self-reported ideology, which we would expect to be stable or slowly moving over time. I find that nearly all of the relationships discussed below prove to be stable, so in most cases I simply discuss the results of a pooled analysis of all four surveys (n=7218).[9] All estimates are provided for the pooled data and for each year analyzed separately in an appendix.

3.1. *The Location of the Primary Axis*

I start by fitting the IRT portion of the model on each NES survey to check for trends in relative values of the question parameters (movement of the principle axis). Table 2 provides the β_j for each question in each year. In contrast to the strong evidence that expressed policy preferences are very unstable at the individual level (Converse 1964, Achen 1975, Zaller 1992), I find almost no change in the relationship between individual polling questions and the primary axis over time. Individual Americans may be shifting their positions from month to month, but I find evidence of aggregate coherence: the primary structural relationship between expressed issue positions has hardly moved at all in the last 20 years.

Despite very little overall change, some issues have become slightly more or less associated with the primary axis over time. Beliefs regarding abortion, affirmative action and spending on public schooling became more strongly correlated with the primary axis while beliefs regarding homosexuality have become less so. Spending issues show only a few trends of substantively interesting magnitude. The correlation between expressed preferences regarding social spending on all issues except crime and spatial position increased between

[9] I include all available respondents (even those who were not asked the self-reported ideology question in 2000) so as to improve the estimation of the primary axis location.

1992 and 1996, perhaps due to the heightened attention to spending issues surrounding the 1994 election and 1995 government shutdown. The isolationism question shows movements consistent with control of the presidency: in every presidential year NES, the sign of the question parameter β_j corresponds to the president's party favoring engagement relative to the out party.[10]

There are large changes in the baseline responses α_j on several issues (see Table 3 for α_j estimates). Average responses moved to the right on defense spending and affirmative action while they moved to the left on the traditional values question, issues related to homosexuality, and the role of women in society. Opinion on welfare moved to the right until 1996 as it polarized and then moved to the left as the issue depolarized.[11] In several cases—welfare between 1992 and 1996, defense spending between 2000 and 2004—shocks appear to both induce a shift in opinion α_j and polarization β_j. This may be the result of the nature of such shocks: they change the responses of the poorly informed rather than more attentive and partisan citizens. Shifts that push moderates and independents toward one side will tend to shift α_j as well as increasing the magnitude of β_j as attention to the issue by the public sharpens differences across the ideological spectrum.

3.2. *Aggregate Issue Positions versus Self-Reported Ideology*

Because of the minimal changes in the location of the primary axis over time, I generally discuss results from pooling all four surveys in the main text and provide results from separate analyses of each survey year in the tables. Pooling has the benefit of putting all four years on the same scale and bringing over 7000 respondents to bear on the questions related to the quality of self-reports. There are a few temporal patterns that are interesting that are mentioned when appropriate. The pooled model allows us to estimate comparable

[10]This is also true in every off-year NES going back to 1988, using only the subsets of the same list of questions that are asked in those years.

[11]These results can also be found in simpler analyses of mean responses.

hierarchical parameters describing the distribution of ideal points in each year, which are plotted in Figure 1. The issue opinions of Americans moved right from 1992-1996 and left from 1996-2004 and appear somewhat more polarized in 1996 and 2004 than in 1992 and 2000. This seems consistent with the timing of the 1994 Republican takeover of Congress, perceptions of 2000 being a low-stakes election, and the relatively polarized 2004 campaign. The variance results for estimated positions are not consistent with the results of analyzing self-reported ideology however: the variance in self-reports follows exactly the opposite pattern of positions.[12] We can understand the greater variance in estimated positions as the result of greater citizen coherence because of greater attention to politics; however, the finding that self-reports do not follow will be revisited below. The mean of self-reports moves right and then left over the same periods as issue positions. But while estimated positions return to their 1992 location by 2000 and are further to the left in 2004, mean self-reports only return to their 1992 position in 2004. Given the shifts in the α_j parameters over time (specifically shifts in responses with regards to women's role in society and gay rights issues) we can explain this result easily: self-reported conservatives are not as "conservative" on the issues in the sample as they used to be.

3.3. *Issue Positions, Spatial Coherence, and Political Information*

A novel feature of the heteroskedastic IRT model employed here is that it allows errors to vary in magnitude across individuals. Substantively, it makes sense that respondents who have less political information and pay less attention will make more ideological "mistakes" than those whose knowledge of politics enables them to form coherent profiles of positions. We can verify this claim by showing that the magnitude of errors made by respondents is greater for lower information/attention groups. To do this, I model the estimated z_i and σ_i hierarchically as draws from distinct normal and inverse-gamma distributions for each

[12]The constant polarization hypothesis is consistent with the variance data for both estimated positions and self-reports.

information/attention group.

Figure 2 shows that the expected results hold in pooled data, they also hold in each survey year taken alone (see Table 4). Consistent with previous literature, the variance of z_i is highly dependent on information/attention, with higher values corresponding to larger variance in position. Similar to the aforementioned findings about the quality of self-reported ideology, I find also that the spatial voting errors made by the less well-informed are larger.[13] Whether this is the result of higher rates of response error on issue questions (Achen 1975) or of less spatial thinking (Converse 1964) is unclear at this stage of the analysis.

3.4. *Demographics and Self-Reported Ideology*

To assess the extent to which demographic variation in self-reported ideology is due to variation between groups along the estimated primary axis, I begin by fitting a pure demographic model that attempts to predict self-reported ideology as a function *only* of demographic variables before adding in estimated position as a predictor. If the coefficients on these variables remain unchanged upon adding the spatial position to the model, this might reflect patterns of bias[14] in self-reporting rather than a proxy effect. However, if adding the spatial position z_i to the model were to cause the demographic coefficients to go partly or completely to zero, this would tell us that the demographic model was predicting self-reports because it proxies for the spatial position. If we observe a sign flip in the demographic coefficients upon adding z_i, this would be a less ambiguous sign of bias.

The left plot in Figure 3 shows the effect on demographic predictors of self-reported ideology of adding estimated spatial position to the model and demographic variables related to age, region, sex, race, and education in the pooled data set (Table 5 breaks out each survey year). While there is general attenuation of the demographic coefficients upon inclusion of

[13]The magnitudes of the differences are underestimated because the estimates of σ_i are biased towards 1 because the small number of questions per person and response categories per question limits the number and size of the spatial "mistakes" an individual can make (Lauderdale 2007).

[14]Or a non-spatial component of ideology, depending on one's perspective.

the summary issue position, most of the coefficients are similar across both models. Men say they are more conservative than women on average. Most of this difference disappears when estimated spatial position is introduced to the model, suggesting that part of the difference in self-reported ideology between men and women is due to different positions but that some is due to non-spatial factors. Examination of the trends by year shows that the raw Male-Female difference in self-reports doubled between 1992 and 1996 while the part of that difference not explicable by actual differences in policy went from zero to about half of the total gap. This result suggests that the gender gap widened in this period; however, it did so primarily as a result of shifts in self-perceptions rather than shifts in issue positions.

Black and hispanic respondents are to the left of white respondents on average; however, it appears that this is reflective of different positions because the coefficients on these variables go to zero (for hispanics, past zero) upon the inclusion of spatial position. In contrast, lower education groups appear to self-report themselves as more conservative in a manner that is almost entirely orthogonal to their estimated positions, perhaps reflecting a simple bias in response due to the perceived relative undesirability of the label "liberal" to the label "conservative". Given an identical spatial position, high school educated individuals self-report their ideology more conservatively than the baseline college educated group.

As we might expect, younger voters self-report as more liberal across all years, even after adjusting for their positions, though the magnitude of this difference has declined from 1992-2004, suggesting a cohort effect that is largely rhetorical. There are no significant differences between the 45-65 and the 65+ groups in any survey year or in the pooled analysis. There is only modest evidence that there are regional biases in self-reports, respondents from the South and Central region of the country self-report as more conservative than those from the Northeast and West; however, the Central region effect is almost entirely explicable by differences in issue positions. The southern tendency to self-report as conservative remains even after including issue positions in the model, suggesting that there may be a regional bias. This is a slight concern for analyses like Clinton (2006) that aggregate self-reports

geographically to estimate local political preferences. Fortunately, these differences appear to mostly accentuate real differences in opinion.

More generally, the pattern of coefficients in Figure 3 is nearly the same in all years. With the same questions being asked and—as we saw in the previous section—the principal axis changing little, the relationship between demographics, spatial position, and self-reported ideology is the same across these 12 years of American history. We can be fairly confident that the observed coefficients reflect real, persistent differences in self-reported ideology.

The right plot in Figure 3 shows the results of the ordered probit model for self-reported ideology as a function of estimated position z, demographic variables, and interactions between position and demographics. I have excluded the age and region variables from these analyses because their interactive effects were found to be negligible. The z coefficient tells us the relationship between the estimated spatial position and the latent self-reported position for the baseline group: white, college-educated women. The coefficients for the demographic variables are then offsets from the zero-point of the latent scale: biases in self-reported ideology relative to the baseline group. The coefficients on the interactions of demographic variables with the estimated position correspond to the sensitivity of self-reported ideology to changes in spatial position relative to the baseline group. To get the z_i coefficient for the particular group of interest, one would simply add the coefficient for z to the coefficient for the relevant interaction term. Groups with negative interaction coefficients are less responsive to their spatial positions when they report their ideology.

Since none of the interactions are negative and greater in magnitude than the z coefficient, Figure 3 indicates that all groups are responsive to spatial position in answering the self-reported ideology questions, but to varying degrees. In general we see evidence that groups that are less in touch with the national political dialogue—whether by virtue of limited education or more limited or disjoint social networks—are less responsive to their position on the collectively defined primary axis. Men have historically been more attentive to politics than women (Campbell et al. 1960, Converse 1964, Delli Carpini & Keeter 1996), have

larger social networks (Zheng, Salganik & Gelman 2006), and they have higher sensitivity to estimated position in self-reporting their ideology.

Blacks and Hispanics are both less responsive to their issue positions in self-reporting than whites, however their trajectories move in opposite directions. In 1992, both groups were similarly informative about their issue positions. Since then, blacks have become less informative while Hispanics have become more, to the point that the latter are no longer distinguishable from whites. Since the standard of comparison is the political axis that explains the most variation in the issue positions of the general population, the fact that blacks do not self-report ideologies that are very informative about their issue positions may reflect a political discourse that is disjoint from that of whites (Harris-Lacewell 2004). Separately analyzing the black sub-sample yields question parameters for issues that are very similar to those of the full sample, but almost no relationship between the estimated positions and self-reports. There is no upward trend in responsiveness as a function of political information or attention, suggesting that blacks are using the terms liberal and conservative in different ways from whites—and from the national political discourse—even though their issue positions largely fit together the same way. Following this line of argument, the convergence of Hispanic responsiveness to that of whites may reflect the process of assimilation into the American political mainstream.

While women may be less attuned to the terms of the national political discourse because they are less attentive to politics in general, and blacks may be engaged in a distinct political discourse with different terminology, those with less education face a simple informational hurdle. This is likely due to a lack of awareness of the location of the political axis, which must be learned from observation of politics. The common theme across these results regarding the responsiveness of demographic groups is that one must be attentive to the national political discourse to self-report one's ideology in a way that reflects the terminology of that discourse. If this is the right explanation for the results, we should observe dramatically different responsiveness of self-reports to issue position as a function of political information

and attention, differences which the demographic variables are primarily proxies for.

3.5. *Information, Attention and Self-Reported Ideology*

We can directly assess the theory that attention to politics and being connected to the political system is a pre-requisite to self-reporting in a way that reflects the primary axis by modeling self-reported ideology as a function of estimated position, political information, and their interactions. A popular approach to including political information in models is to construct an information scale using quiz questions on the NES (e.g. naming congressional candidates), unfortunately few of these questions are asked consistently across the range of survey years. Instead I use two simpler measures, interviewer-assessed political information and self-reported political attention. The first is potentially endogenous because the perceived coherence of respondent answers must surely influence the interviewer's assessment. The second is noisy because it is a self-report and thus the very caveats this paper is concerned about with respect to self-reported ideology may apply (i.e. differential perceptions of the meaning of the categories). If both measures yield the same substantive results however, we can be a bit more confident that being attuned to the political system is really necessary to accurately self-report ideology.

There is a consistent bias towards self-reporting as a conservative. The positions z_i are centered at zero, so the intercepts in Figure 4 indicate that the median individual in the position distribution will self-report to the right of the center of the moderate category. The intercept terms vary from 0.10 to 0.25 and the cutpoint between moderate and leans conservative is fixed at 0.5, so the modal response from the median individual is still moderate, however the distribution of responses is shifted to the right. This is true for all information and attention groups, though the bias is slightly weaker for the more informed/attentive groups.

The interaction coefficients in Figure 4 (and Table 6) indicate a very strong effect of

information: increasing information yields monotonically increasing responsiveness to position. The lowest information group self-reports in a way that is entirely unresponsive to their estimated position. This effect is much stronger than any of the relationships in Figure 3, even the education effects. Doing a similar analysis to that for the interviewer assessment, I find that the most attentive citizens have twice as strong a relationship between their self-reported ideology and their estimated position as the least attentive citizens (Figure 4 and Table 7). Given that attention is itself self-reported and may induce some lying to comply with normative standards of behavior that would reduce the size of the observed effects, these are very strong results. Moreover, they have serious consequences for analysis of survey data. Any study where both political information and self-reported ideology are used as independent variables must take into account this strong interaction. The predictive power of self-reported ideology on other variables will be highly dependent on the political information of the respondent.

It is evident from examining plots of mean posterior estimated position against self-reported ideology (Figure 5)that part—but not all—of the explanation for why the relationship between self-reports and estimated spatial positions is weaker in the lower information and attention groups is that those groups have lower variance in their estimated spatial positions. If the low-information group were to be as responsive to their position in their self-reports as the high-information group they would have to self-report much less variant ideologies as well. We would then expect to see that the standard deviation of self-reported ideology is also much lower in the low-information group than in the highest information group however it is only barely lower in the data. Thus, part of the reason for the weaker responsiveness of the low information respondents is that they are using too much of the self-reporting scale given their generally moderate positions. This might explain why the variance of self-reported ideology does not increase in the years that the variance of issue positions increases: if more attention leads to more accurate reporting, more accurate self-reporting could still lead to less variation in self-reported ideology even if position variation

was increased by greater attention.

4. CONCLUSION

The variability of the relationship between issue positions and self-reported ideology is important to both the methodology of survey analysis and to our understanding of how individuals conceptualize political disagreement. Methodologically, self-reported ideology is problematic as a measure of spatial position. Given the same issue positions, men, southerners, and those without college educations report their ideology to the right of women, non-southerners and those with college educations. The self-reports of women, blacks and those with less than a college education are less responsive to changes in issue positions. It seems likely that black respondents are using different criteria in answering the self-reported ideology question, a result congruous with recent literature showing that blacks conceptualize politics differently from whites. These findings should seriously caution the analyst of public opinion: significant biases might result from analyses that casually use self-reported ideology, political information and race as regressors without addressing their interactive effects and unclear causality.

The results here are consistent with a range of views regarding the relative "noisiness" of self-reported issue positions and ideology. While the perspective taken here is that the weakness of the relationship between self-reported ideology and issue positions is due to uncertainty about the meaning of the labels in terms of issue positions, the data are not inconsistent with a model where low-information respondents have true ideologies that they are accurately reporting, but which they do not know how to map into issue positions. Indeed, many authors have assumed that these self-reports represents a fundamental outlook. In this view, some groups are better than others at figuring out what their issue positions are supposed to be on the basis of their ideology, leading to the observed closer relationship between the two among the well-informed.

While it may be that there is a fundamental outlook that acts in this way, there is little reason to believe that the self-reported ideology questions are measuring it.[15] The claim that self-reported ideology is actually measuring a fundamental outlook with any precision is particularly hard to reconcile with the relative non-response rates in the data. As observed by previous authors (Converse & Markus 1979), non-response is much higher on the ideology question than on any issue question. While the non-response rates on issue positions ranged from 0.01 to 0.15 in the pooled sample, the non-response rate on the ideology question was 0.23. Response rates are higher on straightforward questions (e.g. abortion, more or less spending on particular areas) than on more abstract seven point scales (e.g. government versus individual responsibility for medical care, amount of government services). The ideology self-report question is—for the many reasons outlined at the outset of this paper—more abstract and difficult than any of these single issues because it requires the individual to know where they fit in the broader scheme of American politics across all issues.

Whether issue or ideology questions are more reliable makes an critical difference in a variety of inferential situations. For example, in their assessment of the connection between positions on racial policies and self-reported ideology, Sniderman, Crosby & Howell (2000) begin a discussion of high education respondents by writing, "Now consider liberals and conservatives in a position to understand what their political philosophies require of them." The authors assume that self-reported ideology reveals the respondents *true* outlook across all education groups, and that the hard part for the less well-educated is figuring out what policies are implied by that outlook. The authors take the relative strength of the correlation between ideology and policy position in the high education group as evidence that "liberals and conservatives ... make the choice on the basis of their principles." But the closer relationship between policy position and ideology for higher education groups could as easily

[15]The author is more amenable to a model where it is partisanship that acts as a "true" outlook rather than liberal-conservative ideology. Extending these models to explore the relationship between issue positions, ideology, and partisanship is a worthy topic for future research.

result from more accurate ideology self-reports among the better educated group. Since the authors' central argument is that differences over racial policy positions are fundamentally the result of ideology rather than prejudice, whether political information helps individuals figure out which issue positions are supposed to go with their ideology or whether it helps individuals self-report their ideology more accurately to reflect their issue positions (which might themselves be motivated by prejudice) matters a great deal.

In addition to these methodological lessons, the results here are substantively revealing. Political information acts in multiple ways. Not only must one pay attention to politics to form ideologically "coherent" positions, one must pay attention to politics to be able to report the terminologically correct label for those positions back to an interviewer. Just as the orientation of issues in national politics must be learned, the rhetoric of that politics must be learned: those who do not pay attention will not talk about politics like people who do.

APPENDIX A: MCMC ALGORITHM

I follow the notation in Clinton, Jackman & Rivers (2004) and Cowles (1996) as much as possible. The latent response is y^*_{ij} for the opinion on bill j by respondent i is given by,

$$y^*_{ij} = \boldsymbol{z_i}\boldsymbol{\beta_j} - \alpha_j + \epsilon_{ij} \tag{1}$$

Latent utilities are converted into ordinal voting outcomes using the correspondence $y_{ij} = g \iff \gamma_g > y^*_{ij} \geq \gamma_{g-1}$. On a k category question, the cutpoints $\gamma_0, \gamma_1, \gamma_k$ are fixed to $-\infty, 0, \infty$ for identification. When responses are binary, these are all the cutpoints and the model is the standard probit specification. Whereas the standard Bayesian estimation procedure for the spatial model assumes $\epsilon_{ij} \sim N(0, 1)$, I permit legislators to have varying

error distributions,

$$\epsilon_{ij} \sim N(0, \sigma_i) \tag{2}$$

The prior distribution on the question parameters $\boldsymbol{\beta_j}$ and α_j is $N(\boldsymbol{t_0}, \boldsymbol{T_0})$, where $\boldsymbol{t_0}$ typically is $\mathbf{0}$ and $\boldsymbol{T_0}$ is typically $\kappa \boldsymbol{I_n}$ with κ usually chosen to give a diffuse distribution (25 is commonly used). On the respondent ideal point $\boldsymbol{z_i}$, we put a $N(\boldsymbol{v_i}, \boldsymbol{V_i})$ prior, where typically $\boldsymbol{v_i}$ is $\mathbf{0}$ and $\boldsymbol{V_i}$ is $\boldsymbol{I_n}$. For the respondent errors, a conjugate inverse gamma prior is used, $\sigma_i \sim IG(\frac{c_0}{2}, \frac{d_0}{2})$.

Simulating this model requires several modifications of the Clinton, Jackman & Rivers (2004) Gibbs sampler. First, we must add a step to the sampler to draw σ_i. Second, because of high covariance, to achieve reasonable performance of the sampler, the cutpoints γ_g and the latent data y_{ij} must be drawn jointly. The problem with the naive Gibbs sampler approach is that conditional on the latent data, the cutpoint γ_g can only be in the range between the lowest value of y^*_{ij} where $y_{ij} = g + 1$ and the highest value of y^*_{ij} where $y_{ij} = g - 1$. As the amount of data increases in each category, the size of this window narrows and the sampler takes longer to complete a posterior tour. A Metropolis-Hastings step that solves this problem in ordinal data problems was demonstrated by Cowles (1996) and is employed here.

First, we jointly sample the latent utilities and cutpoints for each question for each respondent, conditional on the question parameters, the respondent positions, and the respondent errors. The Metropolis-Hastings step begins by drawing a proposed vector of cutpoints γ_{new} iteratively from the proposal distribution,

$$g(\gamma_{g,new} | \gamma_{g-1,new}, \gamma_{g,old}, \gamma_{g+1,old}) = N\left(\gamma_{g,old}, \sigma_\gamma\right)$$

truncated to the interval $(\gamma_{g-1,new}, \gamma_{g+1,old})$. The acceptance probability for the new cutpoints is then given by $min(1, R)$ where R is given by the formula,

$$R = \prod_{g=2}^{k} \frac{\Phi\left((\gamma_{g+1,old} - \gamma_{g,old})/\sigma_\gamma\right) - \Phi\left((\gamma_{g-1,new} - \gamma_{g,old})/\sigma_\gamma\right)}{\Phi\left((\gamma_{g+1,new} - \gamma_{g,new})/\sigma_\gamma\right) - \Phi\left((\gamma_{g-1,old} - \gamma_{g,new})/\sigma_\gamma\right)}$$

$$\times \prod_i \prod_j \frac{\Phi\left((\gamma_{y_{ij},new} - \boldsymbol{z_i\beta_j} + \alpha_j)/\sigma_i\right) - \Phi\left((\gamma_{y_{ij}-1,new} - \boldsymbol{z_i\beta_j} + \alpha_j)/\sigma_i\right)}{\Phi\left((\gamma_{y_{ij},old} - \boldsymbol{z_i\beta_j} + \alpha_j)/\sigma_i\right) - \Phi\left((\gamma_{y_{ij}-1,old} - \boldsymbol{z_i\beta_j} + \alpha_j)/\sigma_i\right)}$$

Whether or not the new cutpoints are accepted, new latent data are then drawn conditional on γ_g, for $y_{ij} = g$,

$$g(y^*_{ij}|y_{ij} = g, \gamma, \boldsymbol{z_i}, \boldsymbol{\beta_j}, \alpha_j, \sigma_i) \sim N(\boldsymbol{z_i\beta_j} - \alpha_j, \sigma_i) I(\gamma_g \leq y^*_{ij} < \gamma_{g+1})$$

Second, we sample the question parameters conditional on the latent responses and the respondent parameters. Let $\boldsymbol{Z^*}$ be the $n \times (d+1)$ matrix of respondent positions with typical row $\boldsymbol{z_i}^*$ = ($\boldsymbol{z_i}$, -1) and $\boldsymbol{y_j}^*$ be the $n \times 1$ vector of latent responses for the the jth bill. Also define $\boldsymbol{\Lambda}$ as the inverse variance matrix with $\boldsymbol{\Lambda_{ii}} = 1/\sigma_i^2$ and $\boldsymbol{\Lambda_{ij}} = 0 \; \forall \; i \neq j$.

$$g(\boldsymbol{\beta_j}, \alpha_j | y^*_{ij}, \boldsymbol{z_i}, \sigma_i) \sim$$

$$N\left(\left[\boldsymbol{Z^{*\prime}\Lambda Z^*} + \boldsymbol{T_0}^{-1}\right]^{-1} \left[\boldsymbol{Z^{*\prime}\Lambda y^*} + \boldsymbol{T_0}^{-1}\boldsymbol{t_0}\right], \left[\boldsymbol{Z^{*\prime}\Lambda Z^*} + \boldsymbol{T_0}^{-1}\right]^{-1}\right)$$

The variation in σ_i enters into the sampling of the $\boldsymbol{\beta_j}, \alpha_j$ in a manner familiar from GLS techniques for weighted regression. Where in the canonical model this stage corresponds to sampling a "Bayesian regression" (Clinton, Jackman & Rivers 2004), in this model it corresponds to sampling a "weighted Bayesian regression" with known variance matrix. Substantively, this has the effect that the responses of respondents who make larger errors get lower weight in determining the question parameters. Under the constraint that $\sigma_i = 1 \; \forall \; i$, $\boldsymbol{\Lambda} = \boldsymbol{I_n}$ the sampler becomes the canonical model.

Third, we sample the respondent positions conditional on utilities, the bill parameters

and the respondent errors. Let $w_{ij} = y^*_{ij} + \alpha_j = \boldsymbol{z_i}\boldsymbol{\beta_j} + \epsilon_{ij}$ and $\boldsymbol{B}$ be the $m \times d$ matrix with rows given by $\boldsymbol{\beta_j}$. Since the σ_i are constant across j, the variance matrix is just the identity matrix multiplied by a constant σ_i^2, so it can be removed from the matrix multiplication.[16] The relationship between the posterior variance of z_i and the voting error σ_i is close, as in OLS.[17]

$$g(\boldsymbol{z_i}|y^*_{ij}, \boldsymbol{\beta_j}, \alpha_j, \sigma_i) \sim$$

$$N\left(\left[\frac{1}{\sigma_i^2}\boldsymbol{B}'\boldsymbol{B} + \boldsymbol{V_i}^{-1}\right]^{-1}\left[\frac{1}{\sigma_i^2}\boldsymbol{B}'\boldsymbol{w_j} + \boldsymbol{V_i}^{-1}\boldsymbol{v_i}\right], \left[\frac{1}{\sigma_i^2}\boldsymbol{B}'\boldsymbol{B} + \boldsymbol{V_i}^{-1}\right]^{-1}\right)$$

Finally, we sample the respondent errors conditional on the latent responses, the question parameters and the respondent positions. Let $e_{ij} = y^*_{ij} - \boldsymbol{z_i}\boldsymbol{\beta_j} + \alpha_j$, collapse e_{ij} over j to form $\boldsymbol{e_i}$. Then define the sum of square errors for an individual i's votes as $SSE_i = \boldsymbol{e_i}'\boldsymbol{e_i}$,

$$g(\sigma_i|y^*_{ij}, \boldsymbol{\beta_j}, \alpha_j, \boldsymbol{z_i}) \sim IG(\frac{c_0 + m}{2}, \frac{d_0 + SSE_i}{2})$$

This four stage Gibbs sampler allows us to sample the respondent positions for use in the second stage of the model, which estimates the parameters associated with the ideology self-report. This second stage proceeds in three steps. In the first, as above, a Metropolis-Hastings step is used to simultaneously draw the latent data r_i^*, and the self-report cutpoints θ_k conditional on the positions z_i, the self-reporting parameters λ and κ, the error variance ω^2 and the observed self-report r_i. In the second stage, the self-reporting parameters λ and κ are drawn from a multivariate normal distribution conditional on z_i, r_i^*, and ω^2. Finally, in the last stage, the error variance is drawn from an inverse gamma distribution conditional on z_i, λ, κ, and r_i^*. The mathematical form of these stages are the same as the first, second

[16] As before, under the constraint that $\sigma_i = 1 \; \forall \; i$ this formula is equivalent to that in Clinton, Jackman & Rivers (2004)

[17] This is not an identification problem since posterior variance is a descriptive statistic, not a model parameter.

and fourth described above (i.e. omitting the draw of position) except that there is only one question, so some of the matrix quantities are vectors.

Question ID	Description	Response Range		
VCF0806	Responsibility for Med. Care	Government	1 - 7	Private
VCF0809	Responsibility for Stnd. of Living	Government	1 - 7	Individual
VCF0823	U.S. Concern for World Problems	Isolationism	1 - 2	Engagement
VCF0830	Aid To Minority Groups	Government	1 - 7	Self-Help
VCF0834	Role of Women	Equal	1 - 7	Home
VCF0838	Legal Status of Abortion	Pro-Life	1 - 4	Pro-Choice
VCF0839	Government Services	Decrease	1 - 7	Increase
VCF0843	Military Spending	Decrease	1 - 7	Increase
VCF0853	Need for Traditional Values	More	1 - 5	Less
VCF0867	Affirmative Action	Favor	1 - 2	Oppose
VCF0876	Gay Anti-Discrimination Law	Favor	1 - 2	Oppose
VCF0877	Gay Service in Military	Favor	1 - 2	Oppose
VCF0886	Spending on Poor	Increase	1 - 3	Decrease
VCF0887	Spending on Child Care	Increase	1 - 3	Decrease
VCF0888	Spending on Crime	Increase	1 - 3	Decrease
VCF0890	Spending on Public Schools	Increase	1 - 3	Decrease
VCF0894	Spending on Welfare	Increase	1 - 3	Decrease

Table 1: NES questions used with polarity and ordinal scale.

Question Description	±	1992	1996	2000	2004	Pooled
Resp. for Med. Care	+	0.56	0.66	0.52	0.72	0.59 (0.02)
Resp. for Stnd. of Living	+	0.73	1.02	0.64	0.99	0.81 (0.02)
U.S. Conc. for World	∼	0.13	-0.01	-0.06	0.13	0.05 (0.02)
Aid To Minority Groups	+	0.62	0.77	0.56	0.86	0.70 (0.02)
Role of Women	+	0.30	0.33	0.24	0.30	0.30 (0.02)
Legal Status of Abortion	−	-0.17	-0.18	-0.19	-0.23	-0.18 (0.02)
Government Services	−	-0.79	-0.89	-0.82	-0.81	-0.83 (0.02)
Military Spending	+	0.28	0.26	0.26	0.40	0.25 (0.02)
Need for Trad. Values	−	-0.28	-0.29	-0.23	-0.35	-0.29 (0.02)
Affirmative Action	+	0.50	0.72	0.58	0.84	0.60 (0.03)
Gay Anti-Disc. Law	+	0.50	0.60	0.42	0.37	0.48 (0.02)
Gay Service in Military	+	0.46	0.50	0.29	0.31	0.39 (0.02)
Spend on Poor	+	0.88	0.97	0.94	0.88	0.93 (0.03)
Spend on Child Care	+	0.81	0.91	0.90	0.89	0.86 (0.03)
Spend on Crime	∼	0.20	0.16	0.28	0.08	0.19 (0.02)
Spend on Public Schools	+	0.57	0.70	0.78	0.79	0.68 (0.02)
Spend on Welfare	+	0.67	0.85	0.69	0.78	0.75 (0.02)
Respondents		2485	1714	1807	1212	7218

Table 2: Trends in mean posterior β_j for each question 1992-2004. The ± column is the expected polarity based on the current positions of the two major parties and the directionality of the question scale. The priors on population position make the scales imperfectly comparable over time; however, question parameters appear to be largely constant with a few weak trends (abortion, gay rights, gay military service and school spending). Taking advantage of this stability, the final column shows the results from pooling all four surveys (posterior standard error in parentheses).

Question Description	±	1992		1996		2000		2004		1992-2004	
Parameters		α_j	β_j	α_j	β_j	α_j	β_j	α_j	β_j	α_j	β_j
Resp. for Med. Care	+	0.86	0.56	1.29	0.66	1.10	0.52	1.06	0.72	1.03	0.59
Resp. for Stnd. of Living	+	1.65	0.73	2.07	1.02	1.67	0.64	1.65	0.99	1.72	0.81
U.S. Conc. for World	∼	0.63	0.13	0.63	-0.01	0.62	-0.06	0.88	0.13	0.67	0.05
Aid To Minority Groups	+	1.92	0.62	2.25	0.77	1.98	0.56	1.84	0.86	2.01	0.70
Role of Women	+	-0.08	0.30	0.02	0.33	-0.22	0.24	-0.30	0.30	-0.12	0.30
Legal Status of Abortion	−	1.27	-0.17	1.21	-0.18	1.20	-0.19	1.15	-0.23	1.21	-0.18
Government Services	−	2.18	-0.79	2.16	-0.89	2.37	-0.82	2.31	-0.81	2.24	-0.83
Military Spending	+	1.42	0.28	1.81	0.26	1.94	0.26	1.98	0.40	1.65	0.25
Need for Trad. Values	−	0.09	-0.28	0.00	-0.29	-0.13	-0.23	0.18	-0.35	0.03	-0.29
Affirmative Action	+	0.93	0.50	1.08	0.72	1.13	0.58	1.12	0.84	1.03	0.60
Gay Anti-Disc. Law	+	-0.29	0.50	-0.44	0.60	-0.49	0.42	-0.74	0.37	-0.45	0.48
Gay Service in Military	+	-0.24	0.46	-0.56	0.50	-0.73	0.29	-0.93	0.31	-0.54	0.39
Spend on Poor	+	-0.17	0.88	0.14	0.97	-0.04	0.94	-0.22	0.88	-0.07	0.93
Spend on Child Care	+	0.01	0.81	-0.04	0.91	-0.40	0.90	-0.34	0.89	-0.16	0.86
Spend on Crime	∼	-0.53	0.20	-0.47	0.16	-0.41	0.28	-0.44	0.08	-0.47	0.19
Spend on Public Schools	+	-0.45	0.57	-0.54	-0.88	0.70	-0.90	0.78	0.79	-0.64	0.68
Spend on Welfare	+	1.16	0.67	1.57	0.85	1.16	0.69	0.95	0.78	1.21	0.75
Respondents			2485		1714		1807		1212		7218

Table 3: Question parameter estimates for each survey analyzed separately by survey year as well as pooled over all survey years.

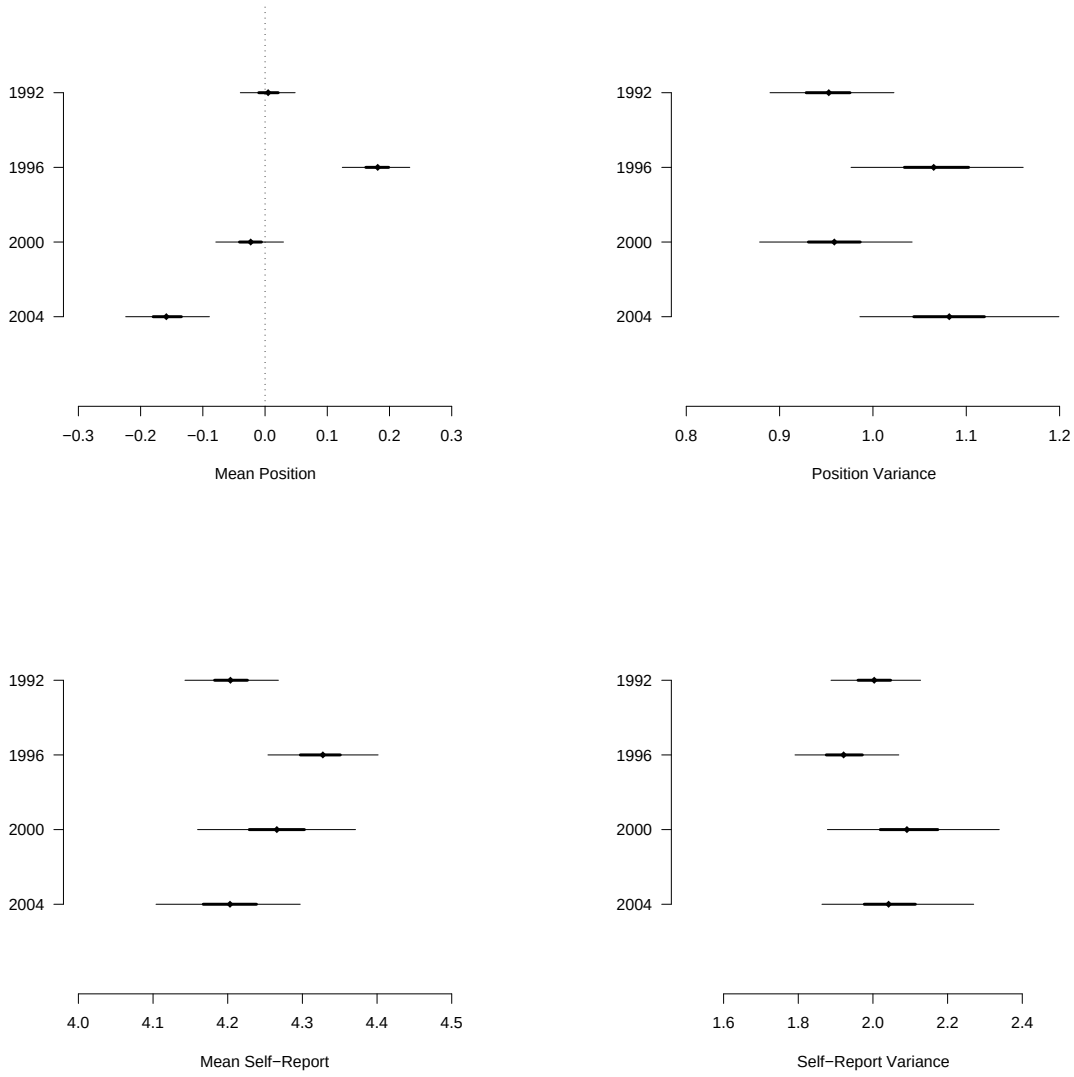

Figure 1: At top left, the mean position of the NES sample based on the pooled estimate is plotted by year. Issue opinions moved right from 1992-1996 and left from 1996-2004. At top right, the variance in positions is plotted by year. It appears that the U.S. electorate was more polarized in the 1996 and 2004 election cycles than in 1992 or 2000, however the greater uncertainty associated with estimating the second moment of a distribution prevents us from ruling out the possibility that there was no change over time. The bottom left and right plots show the analogous analyses based on self-reported ideology. Note that the relative position of the electorate in 1992 and 2004 depends on which measure is used and that the patterns in variance suggest contradictory conclusions about polarization.

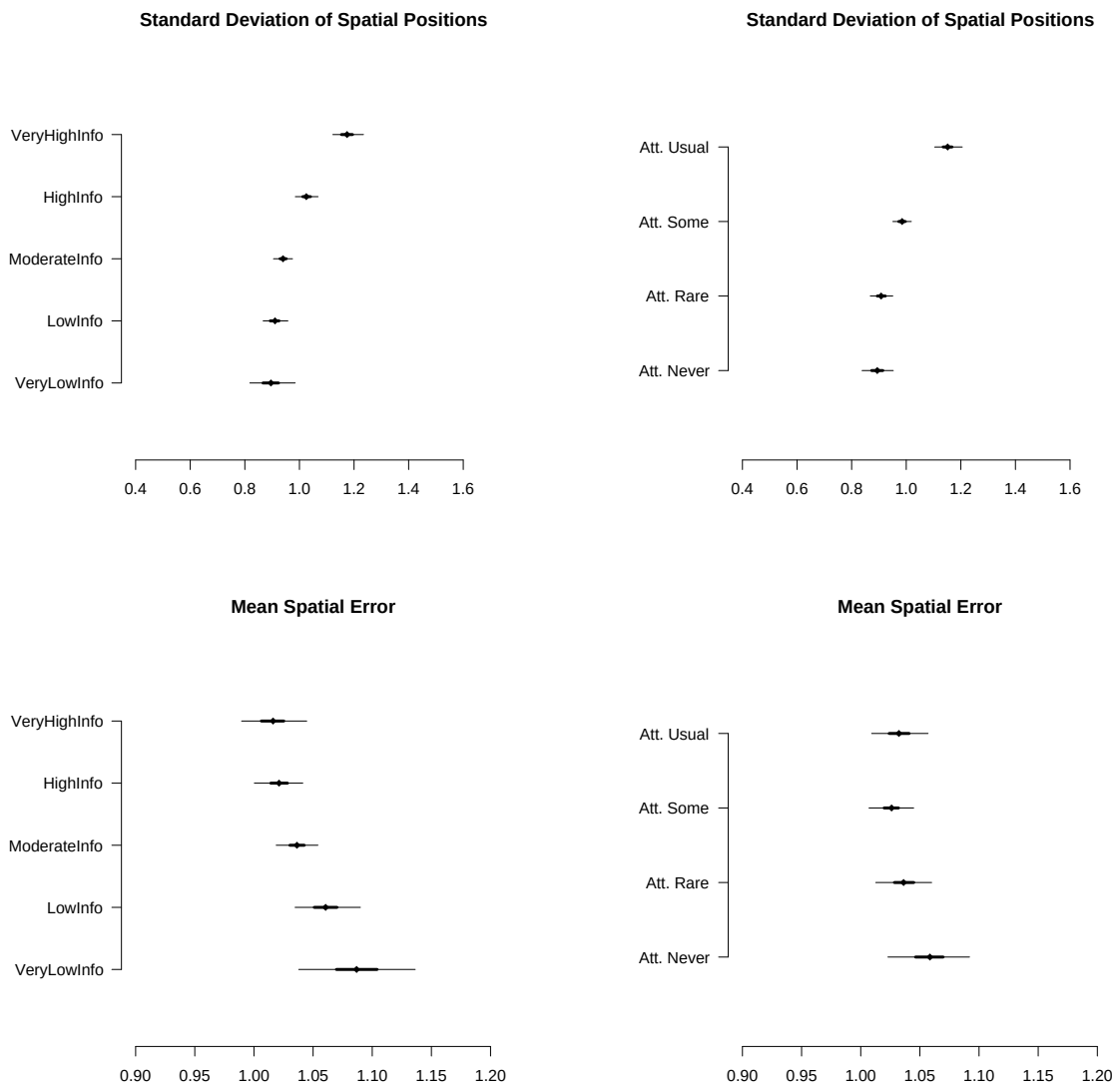

Figure 2: Higher information/attention individuals have greater variance in z_i and lower σ_i. Top pair of plots show the variance of posterior z_i across the population, stratified by information (left) and attention (right). Bottom pair show mean posterior σ_i, which decrease with increasing spatial coherence as information/attention increase.

	$Var(z_i)$					$Mean(\sigma_i)$				
Year	1992	1996	2000	2004	Pooled	1992	1996	2000	2004	Pooled
Very High Info	1.29	1.46	1.24	1.51	1.38	1.02	1.02	1.01	1.02	1.017
	(0.12)	(0.16)	(0.12)	(0.16)	(0.07)	(0.01)	(0.02)	(0.01)	(0.02)	(0.007)
High Info	1.07	1.05	1.09	0.95	1.05	1.01	1.02	1.02	1.03	1.022
	(0.07)	(0.09)	(0.10)	(0.09)	(0.04)	(0.01)	(0.01)	(0.01)	(0.01	(0.005)
Moderate Info	0.91	0.86	0.86	0.81	0.88	1.04	1.03	1.03	1.04	1.036
	(0.05)	(0.07)	(0.07)	(0.08)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.005)
Low Info	0.85	0.77	0.92	0.76	0.83	1.07	1.07	1.05	1.05	1.060
	(0.08)	(0.08)	(0.09)	(0.11)	(0.04)	(0.01)	(0.02)	(0.01)	(0.02)	(0.007)
Very Low Info	0.87	0.82	0.87	0.66	0.80	1.10	1.10	1.08	1.07	1.088
	(0.14)	(0.17)	(0.15)	(0.16)	(0.07)	(0.03)	(0.03)	(0.03)	(0.04)	(0.014)
Never Attention	0.93	0.69	0.82	0.66	0.80	1.08	1.06	1.05	1.05	1.058
	(0.11)	(0.09)	(0.11)	(0.12)	(0.05)	(0.02)	(0.02)	(0.02)	(0.03)	(0.010)
Rare Attention	0.83	0.78	0.87	0.81	0.83	1.04	1.04	1.04	1.04	1.037
	(0.07)	(0.07)	(0.08)	(0.09)	(0.04)	(0.01)	(0.01)	(0.01)	(0.02)	(0.006)
Some Attention	0.94	0.97	0.98	0.98	0.97	1.03	1.03	1.02	1.03	1.026
	(0.06)	(0.07)	(0.07)	(0.09)	(0.04)	(0.01)	(0.01)	(0.01)	(0.01)	(0.005)
Usual Attention	1.30	1.33	1.30	1.34	1.33	1.03	1.03	1.04	1.03	1.032
	(0.09)	(0.12)	(0.12)	(0.13)	(0.06)	(0.01)	(0.01)	(0.01)	(0.01)	(0.006)

Table 4: Left five columns give group variances for estimated positions stratified by information and by attention. Right five columns give group mean spatial response error (σ_i).

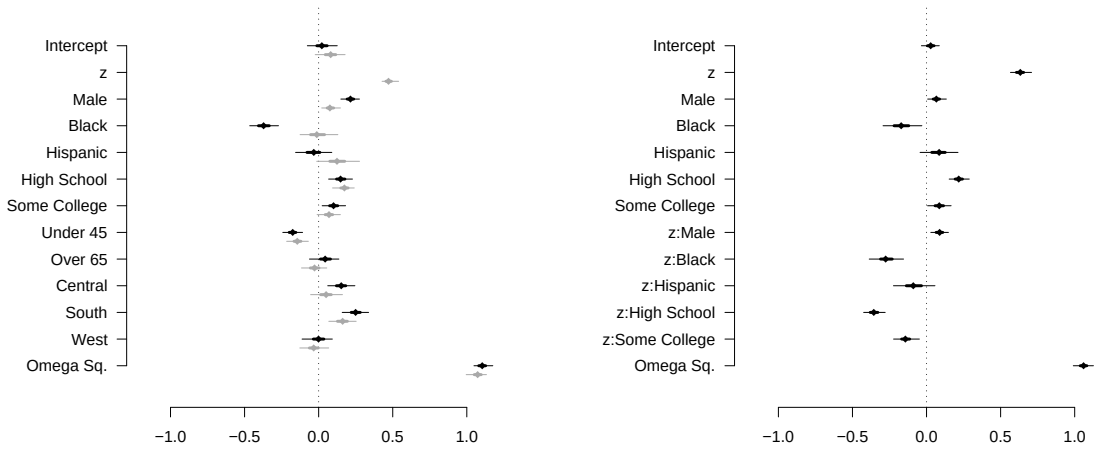

Figure 3: The left plot shows posterior high probability intervals for ordered probit model for self-reported ideology as a function of position z and demographic variables. Results are included for specifications with (grey) and without (black) spatial position. Right plot shows coefficients for a model of demographic characteristics interacted with position to identify which groups provide the most information about issue positions when self-reporting their ideology. The self-reported ideologies of men, whites, and more educated citizens are more predictable on the basis of their issue positions. Coefficient values and errors are provided in Appendix B, Table 5.

Year	1992	1996	2000	2004	Pooled	1992	1996	2000	2004	Pooled	1992	1996	2000	2004	Pooled
Intercept	0.14	0.03	0.02	-0.18	0.03	0.17	0.10	0.07	-0.10	0.07	0.12	-0.02	0.07	-0.07	0.03
	(0.08)	(0.11)	(0.13)	(0.12)	(0.05)	(0.08)	(0.10)	(0.13)	(0.11)	(0.05)	(0.06)	(0.07)	(0.08)	(0.08)	(0.04)
z_i						0.43	0.61	0.39	0.66	0.48	0.60	0.82	0.49	0.80	0.64
						(0.03)	(0.04)	(0.06)	(0.05)	(0.03)	(0.06)	(0.07)	(0.09)	(0.08)	(0.04)
Male	0.14	0.28	0.29	0.23	0.21	0.01	0.13	0.17	0.08	0.09	-0.02	0.13	0.14	0.07	0.07
	(0.05)	(0.06)	(0.08)	(0.08)	(0.03)	(0.05)	(0.06)	(0.09)	(0.08)	(0.03)	(0.05)	(0.06)	(0.08)	(0.08)	(0.03)
Black	-0.48	-0.33	-0.40	-0.20	-0.36	-0.15	0.14	-0.19	0.28	-0.00	-0.24	-0.11	-0.33	0.09	-0.16
	(0.08)	(0.10)	(0.17)	(0.13)	(0.05)	(0.09)	(0.11)	(0.20)	(0.12)	(0.06)	(0.11)	(0.13)	(0.19)	(0.14)	(0.07)
Hispanic	-0.04	-0.04	-0.08	-0.06	-0.05	0.10	0.23	0.03	0.15	0.12	0.02	0.16	0.01	0.12	0.09
	(0.12)	(0.12)	(0.17)	(0.16)	(0.07)	(0.12)	(0.13)	(0.17)	(0.16)	(0.07)	(0.13)	(0.14)	(0.16)	(0.15)	(0.07)
High School	0.10	0.24	0.12	0.18	0.15	0.13	0.27	0.17	0.18	0.17	0.17	0.33	0.18	0.19	0.22
	(0.06)	(0.07)	(0.11)	(0.09)	(0.04)	(0.06)	(0.08)	(0.11)	(0.09)	(0.04)	(0.06)	(0.08)	(0.10)	(0.10)	(0.04)
Some College	0.02	0.18	0.02	0.23	0.10	-0.02	0.16	0.03	0.15	0.07	-0.03	0.19	0.05	0.13	0.08
	(0.07)	(0.08)	(0.10)	(0.09)	(0.04)	(0.07)	(0.08)	(0.10)	(0.09)	(0.04)	(0.07)	(0.08)	(0.10)	(0.09)	(0.04)
Under 45	-0.25	-0.19	-0.09	-0.11	-0.18	-0.16	-0.20	-0.06	-0.05	-0.14					
	(0.06)	(0.07)	(0.10)	(0.08)	(0.03)	(0.06)	(0.07)	(0.10)	(0.08)	(0.04)					
Over 65	-0.02	0.07	0.02	0.12	0.04	-0.07	-0.02	0.02	-0.03	-0.02					
	(0.08)	(0.09)	(0.13)	(0.11)	(0.05)	(0.08)	(0.09)	(0.13)	(0.11)	(0.05)					
Central	0.17	0.11	0.17	0.16	0.15	0.06	0.01	0.02	0.01	0.05					
	(0.08)	(0.10)	(0.13)	(0.12)	(0.05)	(0.07)	(0.09)	(0.13)	(0.11)	(0.05)					
South	0.26	0.18	0.22	0.36	0.25	0.16	0.08	0.14	0.25	0.17					
	(0.07)	(0.09)	(0.11)	(0.11)	(0.04)	(0.07)	(0.09)	(0.12)	(0.11)	(0.05)					
West	0.03	-0.02	-0.04	-0.03	0.00	0.01	-0.08	-0.08	-0.10	-0.03					
	(0.08)	(0.10)	(0.12)	(0.12)	(0.05)	(0.08)	(0.09)	(0.13)	(0.11)	(0.05)					
$z_i \cdot$ Male											0.13	0.01	0.10	0.06	0.08
											(0.05)	(0.06)	(0.09)	(0.08)	(0.03)
$z_i \cdot$ Black											-0.15	-0.34	-0.31	-0.50	-0.27
											(0.09)	(0.12)	(0.18)	(0.14)	(0.06)
$z_i \cdot$ Hispanic											-0.23	-0.06	-0.09	0.02	-0.08
											(0.13)	(0.14)	(0.18)	(0.15)	(0.07)
$z_i \cdot$ High School											-0.39	-0.37	-0.24	-0.35	-0.35
											(0.06)	(0.08)	(0.10)	(0.10)	(0.04)
$z_i \cdot$ Some College											-0.15	-0.19	-0.06	-0.11	-0.14
											(0.07)	(0.08)	(0.10)	(0.09)	(0.04)
ω^2	1.12	1.11	1.13	1.13	1.10	1.07	1.05	1.09	1.03	1.08	1.06	1.04	1.09	1.02	1.06
	(0.05)	(0.06)	(0.09)	(0.07)	(0.03)	(0.05)	(0.06)	(0.08)	(0.08)	(0.04)	(0.05)	(0.06)	(0.09)	(0.07)	(0.04)
Cut 1—2	-1.84	-1.92	-1.83	-1.99	-1.85	-1.98	-2.26	-1.94	-2.27	-2.05	-1.99	-2.26	-1.99	-2.26	-2.10
	(0.06)	(0.07)	(0.10)	(0.08)	(0.04)	(0.06)	(0.10)	(0.10)	(0.09)	(0.05)	(0.06)	(0.13)	(0.10)	(0.11)	(0.06)
Cut 2—3	-0.98	-1.04	-0.95	-1.02	-0.99	-1.02	-1.16	-1.00	-1.13	-1.06	-1.04	-1.15	-0.99	-1.13	-1.07
	(0.03)	(0.04)	(0.05)	(0.04)	(0.02)	(0.03)	(0.04)	(0.05)	(0.05)	(0.02)	(0.03)	(0.05)	(0.05)	(0.05)	(0.02)
Cut 3—4	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5
Cut 4—5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Cut 5—6	1.08	1.05	0.98	0.97	1.03	1.15	1.15	1.02	1.09	1.11	1.16	1.16	1.02	1.10	1.11
	(0.03)	(0.03)	(0.05)	(0.03)	(0.01)	(0.03)	(0.04)	(0.04)	(0.04)	(0.02)	(0.03)	(0.04)	(0.05)	(0.04)	(0.02)
Cut 6—7	2.08	2.16	2.05	2.04	2.08	2.21	2.46	2.17	2.36	2.28	2.26	2.51	2.23	2.41	2.30
	(0.06)	(0.06)	(0.10)	(0.06)	(0.04)	(0.06)	(0.08)	(0.10)	(0.10)	(0.04)	(0.06)	(0.08)	(0.11)	(0.09)	(0.03)

Table 5: Demographic variable analysis of self-reports Left five columns are purely demographic predictors of self-reports. Center five columns add estimate issue position as a predictor. Right five columns add interactions to assess whether certain groups self-report more informatively.

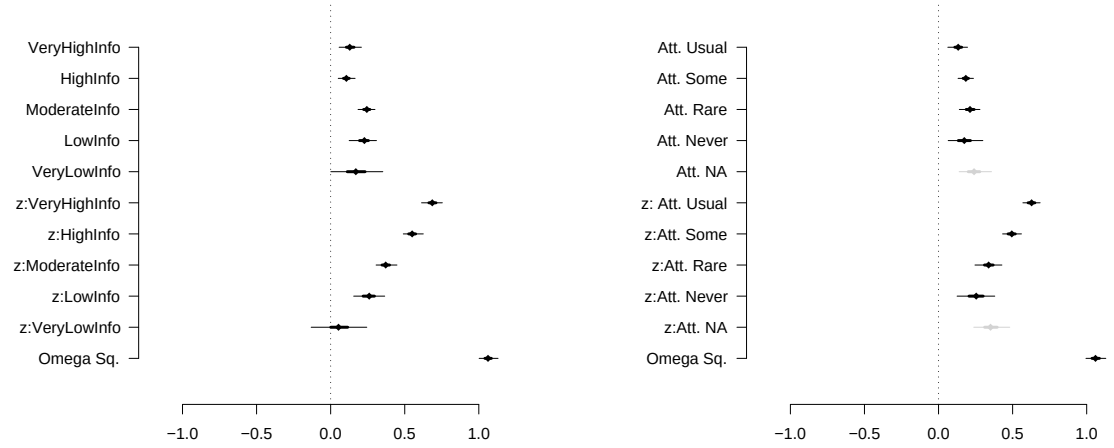

Figure 4: Posterior high probability intervals for ordered probit model for self-reported ideology as a function of position z, political information (left) or political attention (right), and interactions. The trend in the interaction terms demonstrates that there is a very strong dependence of the relationship between position and self-reported ideology as a function of political information/attention. The lowest information/attention groups self-report their ideology with little dependence on z_i while there is a monotonically increasing dependence with increased information.

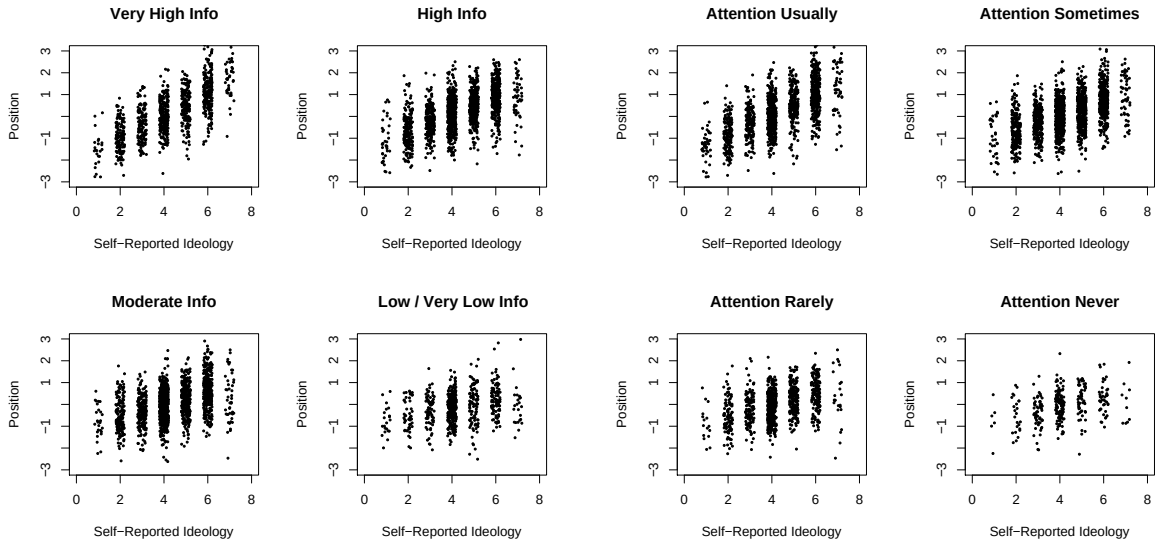

Figure 5: Mean posterior estimated positions as a function of self-reported ideology stratified by political information (left) or political attention (right) for the pooled data. Only higher information/attention groups are providing information about their issue positions when they self-report.

Year	1992	1996	2000	2004	Pooled
Very High Info	0.08	0.20	0.17	0.03	0.13
	(0.06)	(0.08)	(0.10)	(0.08)	(0.04)
High Info	0.14	0.02	0.18	0.08	0.11
	(0.05)	(0.06)	(0.08)	(0.07)	(0.03)
Moderate Info	0.21	0.36	0.20	0.17	0.25
	(0.05)	(0.05)	(0.08)	(0.07)	(0.03)
Low Info	0.27	0.30	0.21	0.16	0.24
	(0.08)	(0.09)	(0.11)	(0.12)	(0.05)
Very Low Info	0.12	0.45	-0.05	0.24	0.19
	(0.18)	(0.24)	(0.23)	(0.24)	(0.11)
$z_i \cdot$ Very High Info	0.66	0.81	0.56	0.81	0.68
	(0.06)	(0.07)	(0.08)	(0.07)	(0.03)
$z_i \cdot$ High Info	0.52	0.69	0.53	0.67	0.55
	(0.05)	(0.06)	(0.08)	(0.07)	(0.03)
$z_i \cdot$ Moderate Info	0.32	0.51	0.32	0.47	0.37
	(0.05)	(0.06)	(0.09)	(0.08)	(0.04)
$z_i \cdot$ Low Info	0.26	0.39	0.11	0.41	0.25
	(0.09)	(0.10)	(0.12)	(0.14)	(0.05)
$z_i \cdot$ Very Low Info	0.02	0.17	0.30	-0.12	0.08
	(0.17)	(0.24)	(0.23)	(0.31)	(0.10)
ω^2	1.05	1.07	1.08	1.02	1.06
	(0.05)	(0.06)	(0.09)	(0.07)	(0.04)

Table 6: Self-reports as a function of estimated issue position and interviewer-assessed political information.

Year	1992	1996	2000	2004	Pooled
Never Attention	0.17	0.16	0.09	0.22	0.15
	(0.11)	(0.11)	(0.14)	(0.14)	(0.05)
Rare Attention	0.33	0.23	0.07	0.12	0.21
	(0.06)	(0.07)	(0.08)	(0.07)	(0.04)
Some Attention	0.13	0.25	0.28	0.11	0.18
	(0.04)	(0.05)	(0.07)	(0.06)	(0.03)
Usual Attention	0.09	0.23	0.21	0.04	0.13
	(0.05)	(0.07)	(0.09)	(0.07)	(0.03)
Attention N/A	0.34	0.29	0.15	-0.55	0.25
	(0.09)	(0.10)	(0.11)	(0.54)	(0.06)
$z_i \cdot$ Never Attention	0.24	0.33	0.24	0.41	0.24
	(0.12)	(0.14)	(0.11)	(0.19)	(0.07)
$z_i \cdot$ Rare Attention	0.31	0.46	0.31	0.46	0.35
	(0.07)	(0.08)	(0.10)	(0.10)	(0.04)
$z_i \cdot$ Some Attention	0.42	0.64	0.45	0.63	0.50
	(0.05)	(0.05)	(0.08)	(0.07)	(0.03)
$z_i \cdot$ Usual Attention	0.58	0.70	0.60	0.80	0.63
	(0.05)	(0.06)	(0.09)	(0.07)	(0.03)
$z_i \cdot$ Attention N/A	0.46	0.47	0.21	-0.18	0.35
	(0.09)	(0.10)	(0.11)	(0.33)	(0.06)
ω^2	1.07	1.05	1.08	1.04	1.07
	(0.05)	(0.06)	(0.09)	(0.07)	(0.04)

Table 7: Self-reports as a function of estimated issue position and self-reported political attention.

REFERENCES

Achen, Christopher H. 1975. "Mass Political Attitudes and the Survey Response." *American Political Science Review* 69(4):1218–1231.

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

Bartels, Larry M. 2007. *Unequal Democracy*. Princeton University Press.

Brady, Henry E. 1985. "The Perils of Survey Research: Inter-Personally Incomparable Responses." *Political Methodology* 11(3):269–292.

Campbell, Angus, Philip E. Converse, Warren E. Miller & Donald E. Stokes. 1960. *The American Voter*. Wiley.

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

Clinton, Joshua D., Simon Jackman & Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98(2):355–370.

Converse, Philip E. 1964. The Nature of Belief Systems in Mass Publics. In *Ideology and Discontent*. New York: Free Press pp. 206–261.

Converse, Philip E. & Gregory B. Markus. 1979. "Plus ca change...: The New CPS Election Study Panel." *American Political Science Review* 73(1):32–49.

Cowles, M. K. 1996. "Accelerating Monte Carlo Markov Chain Convergence for Cumulative-link Generalized Linear Models." *Statistics and Computing* 6:101–110.

Delli Carpini, Michael X. & Scott Keeter. 1996. *What Americans Know About Politics and Why It Matters*. Yale University Press.

Enelow, James M. & Melvin J. Hinich. 1984. *The Spatial Theory of Voting: An Introduction*. Cambridge University Press.

Haberman, Shelby J. 1977. "Maximum Likelihood Estimates in Exponential Response Models." *The Annals of Statistics* 5(5):815–841.

Harris-Lacewell, Melissa V. 2004. *Barbershops, Bibles, and BET: Everyday Talk and Black Political Thought*. Princeton University Press.

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

King, Gary, Christopher J. L. Murray, Joshua A. Salomon & Ajay Tandon. 2004. "Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research." *American Political Science Review* 98(1):191–207.

King, Gary & Jonathan Wand. 2007. "Comparing Incomperable Survey Responses: Evaluating and Selecting Anchoring Vignettes." *Political Analysis* 15(1):46–66.

Knight, Kathleen. 1985. "Ideology in the 1980 Election: Ideological Sophistication Does Matter." *Journal of Politics* 47(3):828–853.

Lauderdale, Benjamin. 2007. "Proxmire and the Golden Fleece: Searching for Maverick Legislators in Spatial Voting Errors." *Midwest Political Science Association Annual Conference* .

Lewis, Jeffrey B. 2001. "Estimating Voter Preference Distributions from Individual-Level Voting Data." *Political Analysis* 9(3):275–297.

Londregan, John. 2000. "Estimating Legislators' Preferred Points." *Political Analysis* 8(1):35–56.

Martin, Andrew D. & Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10:134–153.

Martin, Andrew D. & Kevin M. Quinn. 2007. "R Package "MCMCpack"." .
URL: *http://mcmcpack.wustl.edu/*

Poole, Keith T. 2005. *Spatial Models of Parliamentary Voting*. Cambridge University Press.

Poole, Keith T. & Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting*. Oxford University Press.

Quinn, Kevin M. 2004. "Bayesian Factor Analysis for Mixed Ordinal and Continuous Responses." *Political Analysis* 12:338–353.

Sniderman, P., G. Crosby & W. Howell. 2000. The Politics of Race. In *Racialized Politics*, ed. D. J. Sidanius & L. Bobo. University of Chicago Press.

Stanford University and the University of Michigan. 2005. "The 1948-2004 ANES Cumulative Data File." *The American National Election Studies.* .
URL: *http://www.electionstudies.org/*

Treier, Shawn & Simon Jackman. 2002. "Beyond Factor Analysis: Modern Tools for Social Measurement." *Annual Meeting of the Midwest Political Science Association* .

Treier, Shawn & Sunshine Hillygus. 2005. "The Contours of Policy Attitudes in the Mass Public." *Annual Meeting of the American Political Science Association* .

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. Cambridge University Press.

Zheng, Tian, Matthew J. Salganik & Andrew Gelman. 2006. "How Many People Do You Know in Prison?: Using Overdispersion in Count Data to Estimate Social Structure in Networks." *Journal of the American Statistical Association* 101(474):409–423.