# The Micro-Foundations of Mass Polarization: On-line Appendix

Matthew S. Levendusky
Assistant Professor
Department of Political Science
University of Pennsylvania

December 21, 2008

## Data

In this appendix, I detail the data used in the model. The data on respondent preferences come from the following items: placement on the liberal-conservative self-identification scale (V923509,V940839,V960365), the government services/spending trade-off (V923701,V940940,V960450), the government's role in providing individuals with health care (V923716,V940950,V960479), and the government's role in guaranteeing each citizen a job (V923718,V940930,V960483). The items on government spending ask respondents for their opinion whether government spending should be increased, decreased or stay about the same. Respondents may also volunteer that spending should be "cut out entirely." Given that this response is not offered by the NES, I fold respondents who say "cut out entirely" into the "decreased" category. The items used are whether spending should be increased on: food stamps (V923725,V940822,V960496), welfare (V923726,V940820,V960497), solving the problem of homelessness (V923730,V960501), government assistance to the unemployed (V923816), and poor people (V923817,V960565). I performed a factor analysis on the items; in both years the data supported a two-dimensional interpretation, with the second dimension dominated by the government spending items. Given that all items load on the first dimension, and the theoretical relevance of the government spending items to left-right dimension (Jacoby 1994), a one-dimensional solution seems to be appropriate.

The items are perceptions of the parties and candidates come from their placements on the liberal-conservative self-identification scale (V923514-V923515,V923517-V923518 in 1992, V940841, V940847-V940848 in 1994, and V960369, V960371, V960379, V960380 in 1996), the government services/spending trade-off (V923702-V923705 in 1992, V940931, V940934-V940935 in 1994, and V960453, V960455, V960461, V960462 in 1996) and the guaranteed jobs and a standard of living scale (V923719-V923722 in 1992, V940941, V940944-V940945 in 1994 and V960484-V960485 in 1996).

## Analysis of Alternative Identification Restrictions

In the paper, model identification is achieved by assuming that the aggregate perceptions of the parties are fixed over time. I mentioned in footnote 10 that I have also identified the model using an alternative set of restrictions as a robustness check. Here, I fixed the value of the discrimination parameter on one item (the liberal-conservative self-identification item) to be 1 in both years, and then also constrained the median voter to be 0 in both years (thereby fixing the scale and location of the latent trait, as well as ensuring comparability across years).

In this alternative model, I can test to see whether or not I can distinguish the perceived locations of the parties in 1992 vs. 1996. If you cannot, this suggests that my assumption in the paper is not particularly onerous. Here, the 95% HPD interval on the difference between the Democrats' position in 1996 and 1992 is [-0.025, 0.49], and the corresponding 95% HPD interval for the Republicans is [-0.16,0.33]. In both cases, the HPD interval overlaps 0, so we cannot distinguish 1992 positions from 1996 positions. This should help reassure readers troubled by the identification assumption employed in the paper.

Finally, note that studying change over time, by definition, requires strong assumptions. Something must stay fixed over time in order to define change. In the alternative model, one has to hold other parts of the model fixed: for example, the location of the average voter and the discrimination power of the liberal-conservative item. Is that somehow better? The answer, I suppose is a matter of taste. The point more generally is that this is one assumption, and it does not seem to have overly stringent implications.

## Replication of Table 2 Without Accounting for Measurement Error

Table 1 replicates table 2 from the body of the paper, with one important change. In table 2 in the paper, I reported only changes that you could be distinguished from random noise. Here, I include all changes in the raw data.

[Table 1 about here.]

Here, we see one of the chief advantages of using a method that generates measures of uncertainty for the estimates of the latent trait. In table 1, I have to interpret any observed change as evidence of genuine preference change. As you can see by comparing table 2 to table 1, the results in table 1 suggest nearly everyone in the sample is changing from year to year (note that the not all party-year pairs sum to 100 because some people keep their preferences constant, even in the raw data). But as I discussed in the body of the paper, most of these changes are not genuine changes in that they are not able to be distinguished from random error in my estimates of respondents' preferences. As such, one needs a method that can distinguish true change from random error when studying polarization and related preference changes.

## Additional Validity Checks

Here, I present three additional validity checks for my measure that supplement the correlation analysis given in the body of the paper. First, I present a series of boxplots of my latent issue preference measure against partisanship and ideological self-identification. Next, I discuss the item parameters (the $\beta$ parameters from the model), and finally, I present a series of graphs depicting my measure plotting against the raw data from the model. All three analyses reinforce the conclusion presented in the body of the text: the latent preferences measure developed in the text does, in fact, validly measure respondents' issue preferences.

### Boxplots

It should also be the case that those estimated to be more conservative by my measure are more Republican and are more likely to identify as a conservative. Figure 1 presents boxplots of the posterior mean of the latent trait against the respondent's partisanship and liberal-conservative self-identification in 1992, 1994, and 1996.

[Figure 1 about here.]

The data show the pattern I expected: on average, more Republican and conservative individuals have more conservative issue preferences in all years.[1] That said, however, the relationship between the survey measures and my issue preferences measure is far from perfect at the individual level. For example, many respondents who identify as moderates hold issue positions that are in fact quite liberal or conservative. On average, the liberal/conservative self-identification and party ID scales operate as one would expect (more conservative respondents are more likely to identify as conservatives or Republicans), but for any given respondent, the connection between latent issue preferences and these indicators is more tenuous (Conover and Feldman 1981). Overall, these results demonstrate that my latent preference measure does in fact actually meaningfully measure policy preferences.

### Issue Parameters

As a further validity assessment, I examine the item discrimination parameters. Item parameters that correspond to *ex ante* knowledge about issue positions (i.e., that liberals should favor more government spending) will increase the reader's confidence in my measure. Table 2 gives the relevant posterior means and 95% highest posterior density (HPD) regions.

[Table 2 about here.]

[1]Note that for the liberal-conservative self-placement item, the final group (position 8) represents those who say they "don't know" what their self-identification is.

Recall that the discrimination parameter for each item tells us how strongly the item distinguishes between more or less liberal respondents. First, note that none of the items have a $95\%$ HPD interval that overlaps 0, indicating that all of these items do distinguish between respondents of different ideologies.[2] Those who favor more spending (even if it means higher taxes), favor a greater government role in health care, favor more government guarantees of a basic standard of living, favor addressing the root causes of crime, self-identify as more liberal and favor increased spending on social programs tend to be more liberal, those who favor the reverse are more conservative.

Additionally, table 2 also gives an indication of model fit by calculating the percentage of respondents for which the model correctly predicts the respondent's choice. The fit varies by item, with some items (such as spending on the poor) have excellent fit while others have somewhat poorer fit (for example, whether we should address the root causes of crime or punish criminals more vigorously). This variation reflects the fact that responses to some NES items are determined to a greater extent by the underlying latent issue preferences along this dimension. Rather than dwelling on differences between these items (which would take us too far afield), the key point to take from table 2 is simply that all of the proposed items are related to respondents' issue preferences. This is further evidence in support of my measure's validity.

### Estimates vs. the Raw Data

Figure 2 plots the posterior means from my model against the raw data as a further validity check of my model.

[Figure 2 about here.]

As figure 2 reveals, the results from my model correspond quite nicely to the raw data. Indeed, the correlations (as I report in the paper) all exceed 0.94, suggesting a strong relationship between the two. This is exactly what one would expect: all that's changed is that I've re-scaled the data, but the underlying patterns remain the same.

## Estimation and Inference via Markov Chain Monte Carlo

As mentioned in the text, given the complexity of the model, I adopt a Bayesian approach for estimation and inference. This choice reflects the fact that such a move—particularly the use of Markov Chain Monte Carlo (MCMC) algorithms—greatly simplifies the analysis. In particular, given the dimensionality of the likelihood function, finding its global maximum will be a vexing problem (see below for the likelihood used in this analysis). The model is extremely highly dimensional: there are 597 respondents interviewed in all three waves of the 1992-1996 NES panel data (and each respondent has three latent positions to be estimated, one per panel data wave), multiple candidate

[2] Here, the sign of the discrimination parameters is not particularly informative. All the sign indicates is the scaling of the item. A positive sign indicates that higher responses indicate a more conservative position (so on the self-placement item, a "5" (conservative) is more conservative than a "2" (liberal), and the reverse is true of a negatively signed item). Magnitude, rather than sign, is the information we want to glean from the table.

and party locations, as well as discrimination and threshold parameters for each policy item. Given this, a traditional frequentist assault on the likelihood will be difficult. Even if it succeeds, then one is left with inverting a very large Hessian to invert to obtain standard errors (although bootstrapping might be used to obtain uncertainty estimates).

Ultimately, one could estimate this sort of model with frequentist methods, perhaps using the marginal maximum likelihood methods proposed in Bock and Aitken (1981). However, as Johnson and Albert (1999) discuss, these methods are difficult to apply in practice when the number of dimensions is quite large. Given the availability of software to conduct Bayesian analysis of this sort of problem, a Bayesian approach is somewhat more practical and it is what I employ here.

Before moving to a more detailed discussion of the model, I need to specify the prior distributions for the model parameters. For the item discrimination parameters, I use a vague normal distribution, $\beta_j \sim N(0, 8) \forall j$, and for the latent trait itself, $x_i \sim N(0, 1)$. For the threshold parameters, recall that $\kappa_{j,k} = \sum_{l=1}^{K} \delta_{j,l}$, so I place a prior on $\kappa$ by placing one on $\delta$. Here, $\delta_{j,1} \sim N(0, 100)$, and $\delta_{l,m} \sim exp(2) \forall l \geq 2$.

As a robustness check, I re-estimated the model with somewhat more diffuse priors (e.g., $x_i \sim N(0, 10)$), the substantive conclusions of the model do not change very much: for example, the correlation between the latent trait estimated with this more vague prior and the standard prior discussed above is 0.98.

In the Bayesian setup, one wants to compute the joint posterior density of all the unknown parameters. Here, let $\boldsymbol{\Omega} = \{\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{\beta}, \boldsymbol{\kappa}\}$. Then the posterior density is $p(\boldsymbol{\Omega}|\boldsymbol{Y}, \boldsymbol{W}) = p(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{\beta}, \boldsymbol{\kappa}|\boldsymbol{Y}, \boldsymbol{W})$, where $\boldsymbol{Y}$ is the matrix of policy positions, $\boldsymbol{W}$ is the matrix of candidate/party placements, made by stacking the relevant data across individuals and items. Via Bayes' Rule, this posterior density is proportional to the prior density over the parameters times the likelihood:

$$p(\boldsymbol{\Omega}|\mathbf{Y}, \mathbf{W}) \propto \mathcal{L}\left(\boldsymbol{W}|\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\kappa}\right) \mathcal{L}\left(\boldsymbol{Y}|\boldsymbol{x}, \boldsymbol{\beta}, \boldsymbol{\kappa}\right) p(\boldsymbol{\Omega}) \tag{1}$$

Here, $\mathcal{L}(Y|\boldsymbol{x}, \boldsymbol{\beta}, \boldsymbol{\kappa})$ is the likelihood function for the data $\boldsymbol{Y}$, all other notation follows from the text. To interpret the above, note that the model posterior is a combination of the likelihood from the candidate placement part of the model, the likelihood from the respondents' self-placement data, and the prior 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, 2004) for a review of the MCMC methods and applications in political science; see also Gelman et al. (1995) on Bayesian methods more generally).

Given the model and prior densities described above, there is no conjugacy (normal prior distributions and ordinal data). 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, I initialized the MCMC algorithm with zeros for all parameters except the $\delta$ 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. Likewise, the candidate/party placements on the latent trait were started at the mean response, again re-centered so that 0 is the moderate position. I ran the MCMC algorithm for 75,000 iterations, and discarded the first 25,000 iterations as burn-in phase to ensure that the algorithm had moved away from the arbitrary start values. Using the Gelman-Rubin-Brooks diagnostic in WinBUGS (Brooks and Gelman 1998), a two-chain model with over-dispersed starting values appears to have reached convergence by the end of the burn-in period. After discarding the burn-in iterations, I save every 50th iteration, resulting in 1000 approximately independent draws from the joint posterior density that are then used for the analysis reported in the text.

## References

Bock, R Darrell, and Murray Aitken. 1981. "Marginal Maximum Likelihood Estimation of Model Parameters: Application of an EM Algorithm." *Psychometrika* 46 (4): 443-459.

Brooks, Steven, and Andrew Gelman. 1998. "General Methods for Monitoring Convergence of Iterative Simulations." *Journal of Computational and Graphical Statistics* 7 (4): 434-455.

Conover, Pamela, and Stanley Feldman. 1981. "The Origins and Meaning of Liberal/Conservative Self-Identifications." *American Journal of Political Science* 25 (4): 617-645.

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

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

Jackman, Simon D. 2004. "Bayesian Analysis for Political Research." *Annual Review of Political Science* 7 (1): 483-505.

Jacoby, William. 1994. "Public Attitudes Toward Government Spending." *American Journal of Political Science* 38 (2): 336-361.

Johnson, Valen, and James H. Albert. 1999. *Ordinal Data Modeling*. New York: Springer-Verlag.

Neal, Radford. 2003. "Slice Sampling (with Discussion)." *Annals of Statistics* 31 (3): 705-767.

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

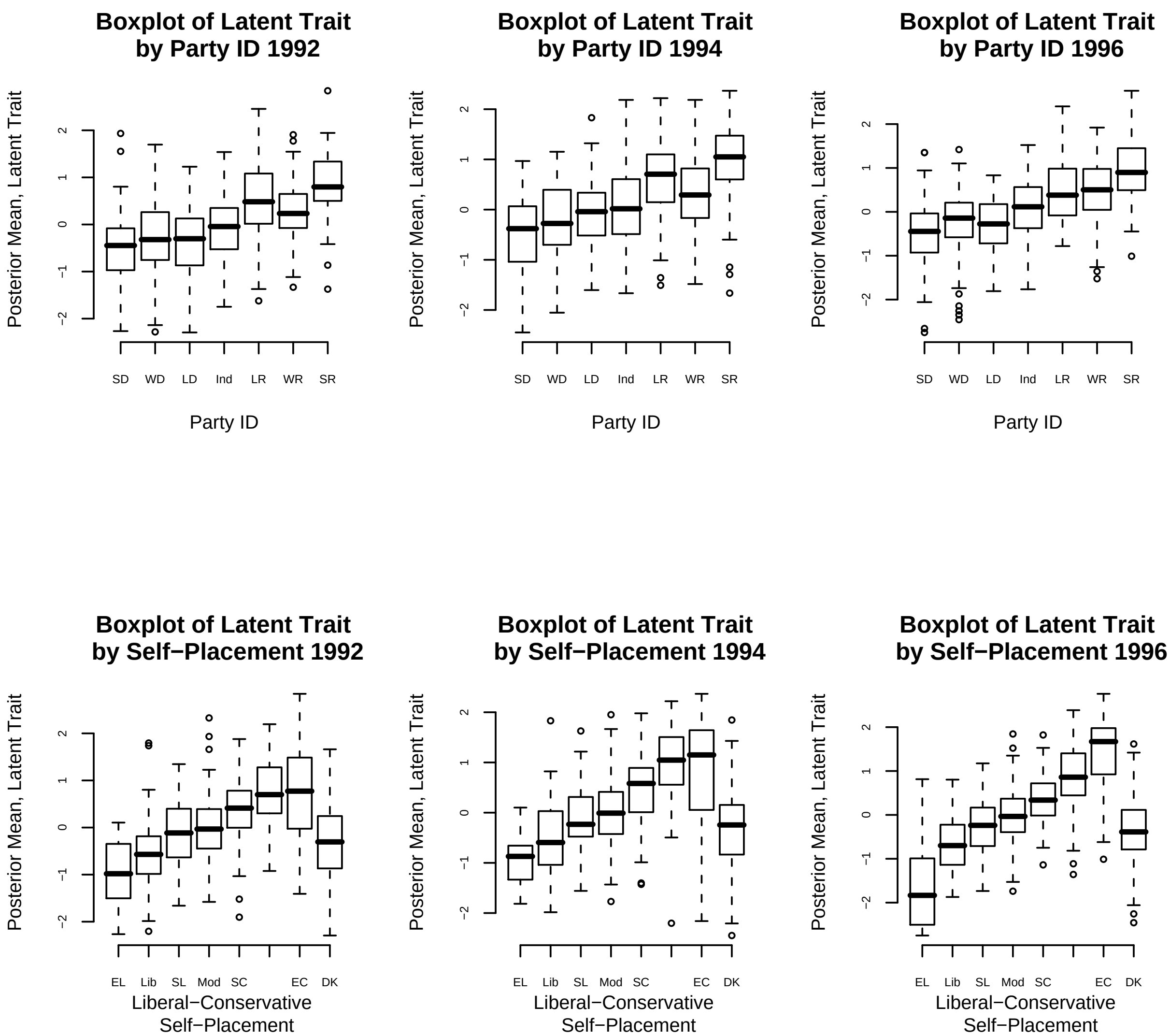


Figure 1: Boxplot of the posterior means of the latent trait by the respondent's partisanship (left panel) and liberal-conservative self-identification (right panel). For the liberal-conservative self-identification scale, positions 1-7 correspond to the NES scale positions, position 8 is for those who say "don't know" to the item.

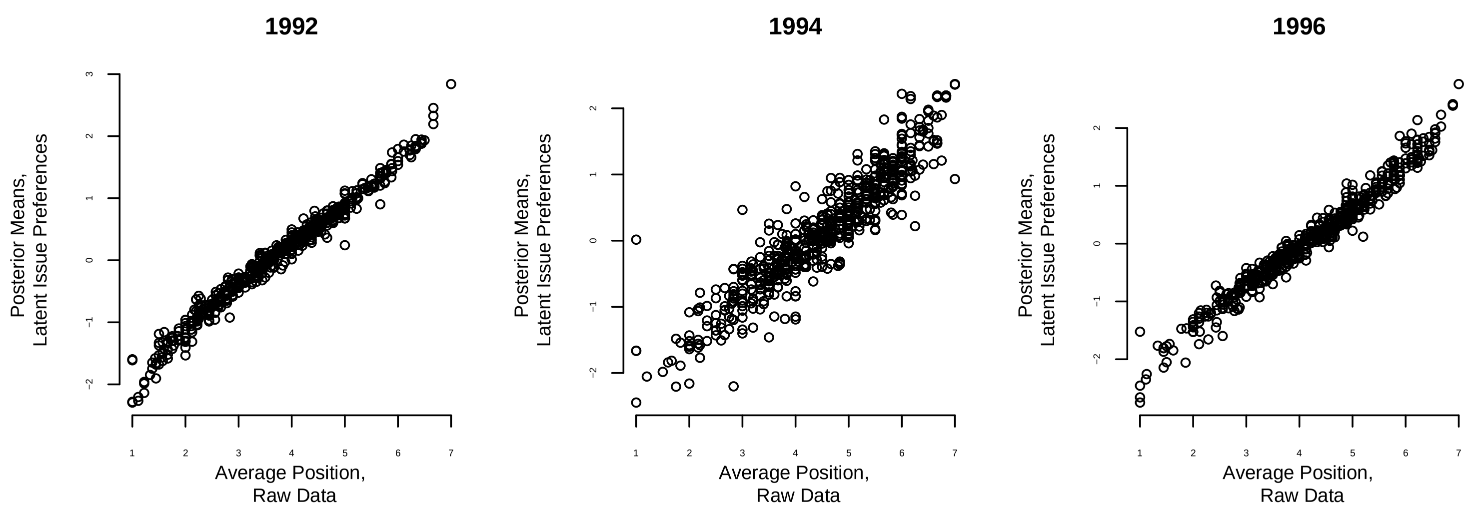


Figure 2: Plot of the raw data against the posterior means of the latent trait in each year.

| | 1992-1994 | 1994-1996 | 1992-1996 |
|---|---|---|---|
| Democrats: | | | |
| Liberal Change | 15 | 66 | 23 |
| Conservative Change | 85 | 32 | 76 |
| Republicans: | | | |
| Liberal Change | 15 | 66 | 25 |
| Conservative Change | 83 | 32 | 70 |

Table 1: Amount of change in the raw data, by party and year. Cell entries are percentages of each party making a given change in a given year.

| | Estimates | PCP |
|---|---|---|
| **7 Response Options** | | |
| Liberal-Conservative Self ID, 1992 | 1.36 | 32.16 |
| | [ 1.26 , 1.46 ] | [ 0.29 , 0.36 ] |
| Government Services/Spending, 1992 | -1.17 | 33.8 |
| | [ -1.26 , -1.08 ] | [ 0.31 , 0.37 ] |
| Health Care, 1992 | 1.05 | 23.84 |
| | [ 0.86 , 1.28 ] | [ 0.2 , 0.27 ] |
| Guaranteed Jobs, 1992 | 1.24 | 26.7 |
| | [ 1.15 , 1.34 ] | [ 0.24 , 0.3 ] |
| Liberal-Conservative Self ID, 1994 | 0.78 | 33.23 |
| | [ 0.71 , 0.85 ] | [ 0.3 , 0.37 ] |
| Guaranteed Jobs, 1994 | 1.33 | 31.67 |
| | [ 1.22 , 1.43 ] | [ 0.28 , 0.35 ] |
| Government Services/Spending, 1994 | -1.24 | 33.29 |
| | [ -1.35 , -1.14 ] | [ 0.3 , 0.37 ] |
| Health Care, 1994 | 1.61 | 31.68 |
| | [ 1.32 , 1.92 ] | [ 0.28 , 0.36 ] |
| Liberal-Conservative Self ID 1996 | 1.46 | 36.42 |
| | [ 1.36 , 1.56 ] | [ 0.33 , 0.4 ] |
| Government Services-Spending 1996 | -1.46 | 39.56 |
| | [ -1.56 , -1.36 ] | [ 0.36 , 0.43 ] |
| Health Care 1996 | 1.4 | 28.28 |
| | [ 1.15 , 1.67 ] | [ 0.25 , 0.32 ] |
| Guaranteed Jobs 1996 | 1.87 | 36.34 |
| | [ 1.57 , 2.21 ] | [ 0.33 , 0.4 ] |
| Crime 1996 | 0.78 | 22.39 |
| | [ 0.58 , 0.99 ] | [ 0.2 , 0.25 ] |
| **3 Response Options** | | |
| Food Stamp Spending 1992 | 2 | 65.12 |
| | [ 1.63 , 2.4 ] | [ 0.62 , 0.68 ] |
| Welfare Spending 1992 | 2.16 | 65.55 |
| | [ 1.76 , 2.59 ] | [ 0.62 , 0.69 ] |
| Spending on the Homeless 1992 | 1.77 | 73.64 |
| | [ 1.43 , 2.16 ] | [ 0.71 , 0.76 ] |
| Spending on the Unemployed 1992 | 1.72 | 64.86 |
| | [ 1.41 , 2.08 ] | [ 0.62 , 0.68 ] |
| Spending on the Poor 1992 | 2.08 | 69.84 |
| | [ 1.68 , 2.5 ] | [ 0.67 , 0.73 ] |
| Welfare Spending, 1994 | 1.91 | 69.63 |
| | [ 1.5 , 2.36 ] | [ 0.66 , 0.73 ] |
| Spending on Food Stamps, 1994 | 1.61 | 67 |
| | [ 1.29 , 1.99 ] | [ 0.64 , 0.7 ] |
| Spending on Food Stamps 1996 | 1.6 | 64.1 |
| | [ 1.3 , 1.95 ] | [ 0.61 , 0.67 ] |
| Spending on Welfare 1996 | 2.07 | 70.82 |
| | [ 1.67 , 2.51 ] | [ 0.68 , 0.74 ] |
| Spending on the Homeless 1996 | 1.49 | 63.85 |
| | [ 1.19 , 1.82 ] | [ 0.61 , 0.67 ] |
| Spending on the Poor 1996 | 2.2 | 68.96 |
| | [ 1.81 , 2.67 ] | [ 0.66 , 0.72 ] |

Table 2: The posterior means and 95 percent HPD intervals for the item discrimination parameters. Items are separated by the number of response options given in the NES. PCP gives the percentage of responses correctly predicted by the model.