Four Essays on the Empirical Analysis of Political Ideology

Thesis by
Peter W. Foley

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2013

(Defended May 7, 2013)

Acknowledgements

I am particularly thankful to Maggie and Helen for keeping me grounded and happy during the course of this research.

I thank Mike Alvarez for his near decade of advice and guidance, Catherine Holcomb, Tim Armstrong and Helen Foley for assistance in running experiments, Gloria Bain for administrative support, and Antonio Rangel for useful discussions on the design of the priming experiment. The advice and friendship of the Adolphs and Rangel labs, and others in the social science department have made me a happier person and improved the quality of this and other work.

Four Essays on the Empirical Analysis of Political Ideology

by

Peter W. Foley

In Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

Abstract

This thesis examines four distinct facets and methods for understanding political ideology, and so it includes four distinct chapters with only moderate connections between them. Chapter 2 examines how reactions to emotional stimuli vary with political opinion, and how the stimuli can produce changes in an individual's political preferences. Chapter 3 examines the connection between self-reported fear and item nonresponse on surveys. Chapter 4 examines the connection between political and moral consistency with low-dimensional ideology, and Chapter 5 develops a technique for estimating ideal points and salience in a low-dimensional ideological space.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Following this introduction, this thesis is broken into four distinct chapters. Chapter 2 examines how reactions to emotional stimuli vary with political opinion, and how the stimuli can produce changes in an individual's political preferences. Ralph Adolphs and R. Michael Alvarez wrote portions of the first chapter, but I performed the experiment and analyses.

Chapter 3 examines the connection between self-reported fear and item nonresponse on surveys. This project is a collaboration with Laura Loesch, R. Michael Alvarez, and Lonna Atkeson. Preliminary analyses, modeling ideas, and introductory text were produced by my collaborators, but I performed all the analyses and results reported here.

Chapter 4 examines the connection between political and moral consistency with low-dimensional ideology, and Chapter 5 develops a technique for estimating ideal points and salience in a low-dimensional ideological space. Both of these chapters are entirely my own.

Chapter 2

The effect of emotional priming on survey responses

Decades of research in political science have assumed that experiences are filtered and processed by a variety of psychological mechanisms to form political attitudes and determine behavior (Campbell *et al.*, 1960; Fiorina, 1978; Jennings and Niemi, 2009). This general causal framework is applied to studies of partisanship and ideology (Campbell *et al.*, 1960), policy opinions (Alvarez and Brehm, 1998, 2002), or participation and candidate choice (Alvarez and Nagler, 1995, 1998; Alvarez *et al.*, 2000; Brady, 1985; Fiorina, 1978; Verba *et al.*, 1995). Only quite recently has research begun to explore the extent to which political attitudes and behavior, and the related psychological mechanisms, might have underlying genetic bases or might be associated with physiological factors (Alford *et al.*, 2005, 2009; Fowler and Dawes, 2008; Mattes *et al.*, 2010; Spezio *et al.*, 2008).

In an important report, Oxley, Smith, et al. (Oxley *et al.*, 2008, hereafter OS) published findings that individuals who favored more socially protective policies like opposition to same-sex marriage and support for the Patriot Act had stronger physiological responses to surprising and threatening stimuli. These findings supported the interpretation that social policy attitudes are connected to the mechanisms that control reactions to emotional stimuli. However, while the OS report received significant attention, it has not been subjected to replication or extension.

As part of a larger research project, we replicated the OS protocol (stimuli and methods) but obtained completely different results. We also used an improved protocol that used a larger set of well-characterized images, randomized the stimulus order, and used a metric for skin conductance response that is less influenced by earlier responses. OS argued, on the basis of their original findings, that conservative subjects tended to react more strongly to threatening stimuli than did liberal subjects, but our data do not support this conclusion. In this study, liberal subjects had significantly stronger skin conductance responses across all types of emotional stimuli, and there was little variation across stimulus types in the reactions of liberals versus conservatives. We extend the physiological measurements to include respiratory sinus arrhythmia, for which we observe emotional category-specific differences between the groups, which could signal differences in emotion regulation. Finally, the results from the image stimuli are shown to carry over to emotion-inducing video stimuli.

2.1 Results

Subjects came to the lab for two sessions. In the first session, they completed a variety of political and psychological questionnaires and then viewed a series of emotional images and a video. In the second session a few months later, they completed the OS protocol and then viewed a different set of emotional images and a video. The images and videos were classified to induce fear, disgust, or be neutral positive, and subjects were randomly assigned to treatments in each session. Details of the experiment and stimuli are in the Methods section and Supplemental Information section at the end of this chapter.

2.1.1 Replication of OS protocol

To replicate OS as closely as possible, we obtained the exact images and the fixed presentation order used in OS from the authors, and we displayed the images for 20 seconds each with 10 second blank screens between images, just as in the original

experiment. Following the analysis of OS, we took the log of skin conductance and subtracted the mean in the 10 second gap preceding an image from the mean during the 20 second image presentation to get a skin conductance response for each individual to each image. We then averaged this mean response for each individual across the three threatening images and the three non-threatening images.

To parallel OS, we classified subjects as having high and low levels of support based on whether they were above or below the median in their support for 'socially protective policies' measured using the same Wilson-Patterson (WP) battery as OS, which had subjects report whether they agree or disagree with a list of policies. This battery was collected from subjects during the first session, meaning that as in OS, our subjects waited a few months after completing the WP battery before we measured their skin conductance responses to the OS images. We grouped some of the WP answers into a single metric of support for socially protective policies as described in the OS supplemental materials. Fig. 2.1 compares the log skin conductance responses between our high- and low-support groups for both threatening and non-threatening images, and the corresponding results from OS are also displayed for comparison.

In OS, subjects with high support for socially protective policies had stronger reactions to the threatening images than subjects with low support, and on non-threatening images, skin conductances were similar across the two groups. In our sample, the group with high support for protective policies has slightly stronger (not statistically significant) reactions to threatening images, but showed a much stronger reaction to non-threatening images than those with low support. Note that because the populations are different and high and low support for protective policies are defined around the median level of support, our high- and low-support groups are somewhat different from those in OS. OS's subjects were selected for high levels of participation, and so they had very few subjects near the middle of the scale, so most of their subjects would be classified into the same group under a range of cutoffs.

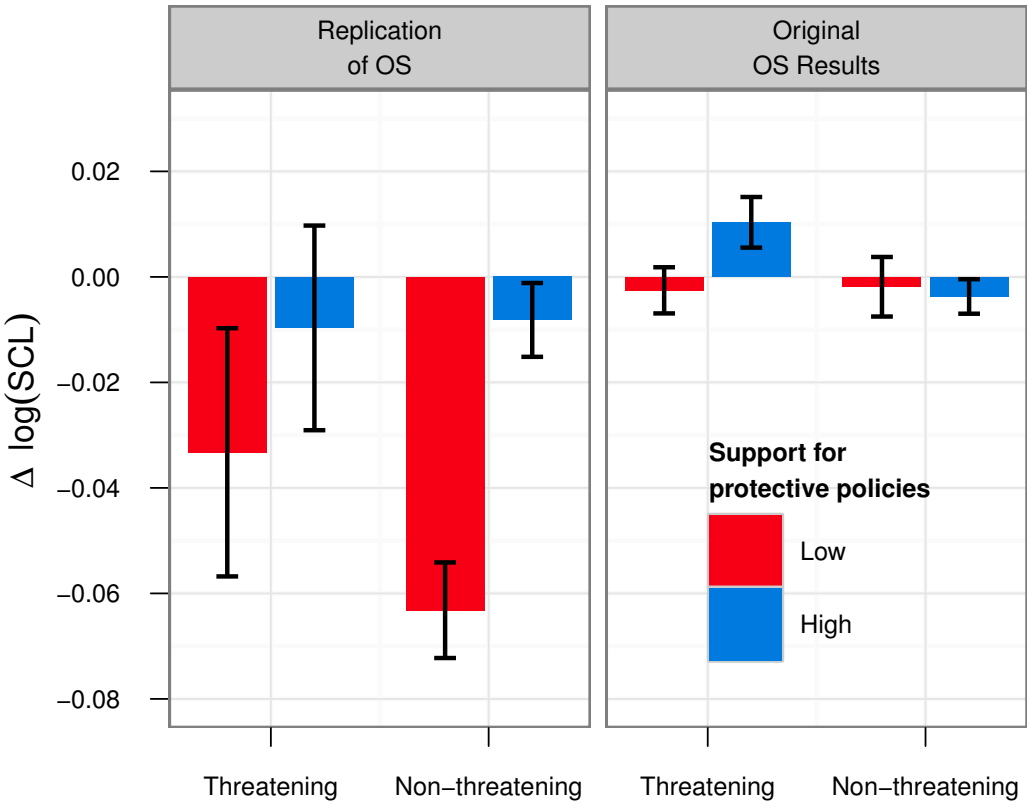


Figure 2.1: Change in log skin conductance upon viewing OS threatening and non-threatening images for groups with low and high support for socially protective policies. The new data was recorded in the second session, and the results from OS are shown for comparison.

Table 2.1: Correlations between responses to threatening and non-threatening OS images using the OS metric

	Spider	Maggots
Maggots	0.235	
Bloody	0.083	0.158

(A) Threatening images

	Rabbit	Fruit
Fruit	0.277	
Happy	-0.055	-0.043

(B) Non-threatening images

2.1.2 Selection of stimuli

Even without the differences shown in Fig. 2.1, there are good reasons to be concerned about the reliability of the OS methods. They used only three 'threatening' and three 'non-threatening' images each out of a total of 25 images shown to subjects, and only one of their images had been previously characterized in terms of emotional content.[1] The threatening images were a spider on a man's face, maggots in a wound, and a man with a bloody face and shirt, but when our subjects saw the images, there was little consistency in individual reactions to images in each set; the cross-correlations in SCR signals using the OS metric are shown in Table 2.1.

2.1.3 Order effects and measurement problems

The ordering of the OS images was also fixed. Though it is easier to find differences between subjects who all observe the same ordering, skin conductance responses to stimuli are characterized by a quick jump upwards a few seconds after stimulus onset followed by a roughly exponential decay, and the time constant of that decay is long enough that the response from one image spills over into the next.

The spillover of reactions between images was especially problematic because of the metric they used for skin conductance response. OS calculated SCR's by

[1]The image of the bloody man in OS is IAPS code 3550 and was characterized by (Mikels *et al.*, 2005) to elicit a very non-specific mixture of emotions.

subtracting the log skin conductance during each image from the log skin conductance during the 10 second before each image, but this metric does not account for the slow decay time and can thus produce negative autocorrelation between successive images. A strong response to one image will result in a jump in skin conductance followed by a slow decline that extends into the next image's time, and this slow decline will bias upwards the skin conductance measured in the next 10 second gap relative to the skin conductance during the next image display, and so the calculated response to the next image will be biased downwards.

Because of the fixed ordering, the same images spilled over into each other for all subjects, and so the errors carry through to the group level. There is no way to know, then, whether reactions to 'threatening' images are really reactions to those images or artifacts of reactions to images that came before. So, while it is valid to say that OS showed a difference between groups, there is no way to interpret that difference.

The phasic and driver metrics for SCR (described in Methods) are less sensitive to order effects than the OS metric, but in our IAPS images where randomization averages away order effects, and where we have enough images for reliable measurements, the three metrics look very similar (Fig. 2.6). For the OS protocol with fixed ordering and only 3 of the images of each type, there are clear differences between the three metrics that highlight the impact of order effects (Fig. 2.5).

For the OS images and fixed order, the OS metric shows conservatives reacting more strongly than liberals to both threatening and non-threatening stimuli, the phasic metric shows conservatives reacting less to threatening images but more to non-threatening ones , and finally, the driver metric shows conservatives reacting less across both stimulus types. The driver metric, which is the most robust to order effects and non-biological noise, is not affected by the design problems of the OS protocol, and so its results for the OS images are very similar to those of our improved design.

2.2 New protocol

Fig. 2.2a shows the reactions of liberal and conservative groups across the three image treatments. We use the driver metric for skin conductance because, compared to the available alternatives, it is robust to order effects, gives a more direct picture of the underlying SNS activation, and is usable over long stimuli like our emotional videos. Mean skin conductance response in every category is higher among those with low support for protective policies. The differences are not statistically significant in any category alone ($p > 0.2$), but they are borderline significant ($p = 0.052$ by two-sided permutation test) when taken together.

The group-level consistency in reactions signals that individual reactions are likely correlated across the treatments, and this is generally the case, as shown in Fig. 2.3a. Reactions to the neutral positive and fear treatments are clearly strongly correlated, but the disgust treatment is somewhat less correlated with the other two. There do not appear to be strong order effects, so, for example, someone who saw the fear condition in the first session and the disgust condition in the second session had reactions similar to those of someone who did the reverse.

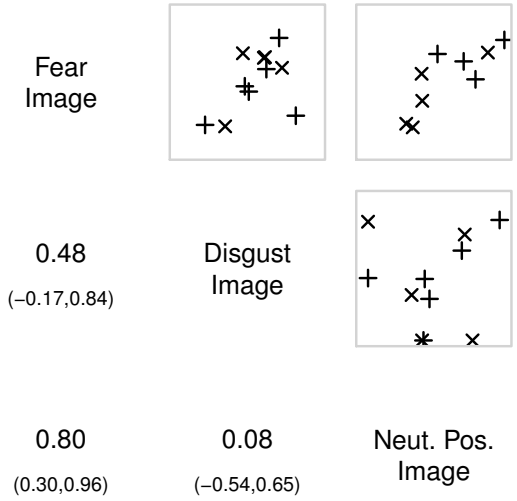
For verification, we also examined subjects' reactions to the emotional videos and found they were very similar to the image reactions. Skin conductance responses during the video (Fig. 2.2b) are very similar to those for IAPS pictures (Fig. 2.2a). Individuals with low support for socially protective policies tend to respond more strongly across all treatments, and as with the images, though the differences are not statistically significant separately ($p > .15$), they are quite significant when viewed across all three treatments ($p = 0.016$ by two-sided permutation test).

Finally, the correlation in skin conductance responses between video treatments (Fig. 2.3b) was similar to what we observed with the images. Individuals who reacted strongly to the fear videos tended to also react strongly to the neutral positive videos, but reactions to disgust videos were less correlated with either of the other two treatments.

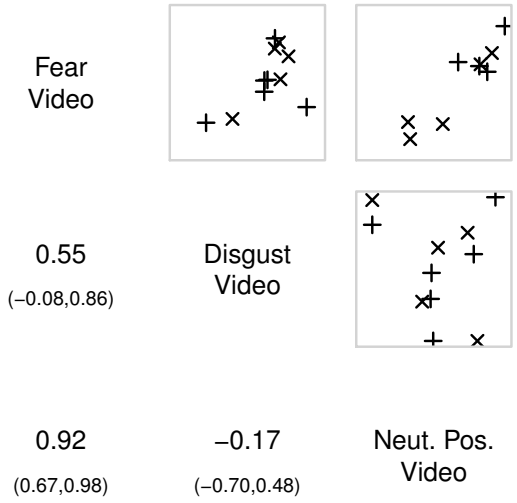
(a) Images

(b) Videos

Figure 2.2: Driver SCR to neutral positive, disgusting, and frightening (a) images and (b) videos across groups with low and high support for socially protective policies. Boxes show the mean and standard errors of responses in each group.



(a) Images



(b) Videos

Figure 2.3: Scatterplots and correlations of reactions to different treatment (a) images and (a) videos across trials. 95% confidence intervals are also shown for the correlations. Subjects who received the row treatment in the first session are marked +, those who saw the column treatment first are marked x.

2.2.1 PNS activation

We also compared groups and treatments in terms of respiratory sinus arrhythmia (RSA), which is generally viewed as measuring activation of the peripheral nervous system (Brownley *et al.*, 2000). PNS activation tends to reduce a person's heart rate, but PNS input to cardiac cells is diminished momentarily during inhalation. This connection results in synchronized variation—the RSA—in respiratory and instantaneous heart rate signals that depends on the level of PNS activation (Brownley *et al.*, 2000). We estimate this synchronization using the wavelet coherence approach from (Keissar *et al.*, 2009) over the high-frequency band (.14-.4 Hz), though we compare the results of a more traditional peak-valley method (Grossman *et al.*, 1990) in the SI.

Previous work has shown a connection between resting RSA and emotions (Oveis *et al.*, 2009), and to reactions and emotion regulation in response to emotional stimuli (Demaree *et al.*, 2004) or during social interactions (Butler *et al.*, 2006), but we do not observe a connection between baseline RSA and social policy preferences (Fig. 2.7). Our baseline RSA is calculated during the pre-stimulus questions in the first session or during the OS images in the second session, and neither show much of a difference between the two groups.

As shown in Fig. 2.4A, RSA differences between social liberals and conservatives do appear following the treatments, but only in the fear condition. In this context, increased RSA can be interpreted as increased emotion regulation, or an effort by the subjects to limit the impact of the emotional stimuli. Individuals in general tend to regulate their emotions in the disgust condition without any observable difference between social liberals and conservatives. Under the fear condition, however, socially liberal subjects appear to increase their emotion regulation while conservatives either do not change or actually decrease regulation. The group difference is highly significant ($p = .015$).

Note that while RSA is a fairly reliable indicator of PNS activation, the step from PNS activation to emotion regulation relies on reverse inference. In this context, though, when the only stimuli are emotional images and videos, there

are no other clear drivers of PNS activation, so interpreting RSA as a measure of emotion regulation is reasonable.

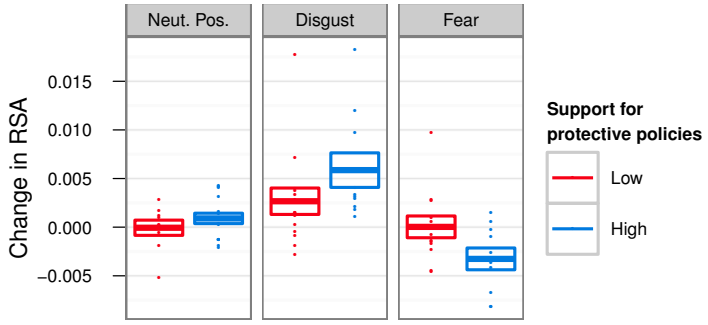
2.2.2 Reactions to video stimuli

The effects of the videos on RSA are shown in Fig. 2.4B and are very similar to what we obtained when subjects reacted to the images. For the RSA metric, since changes in emotional state and regulation are slow, this is not an entirely independent verification of the image results, but the videos are all long enough that these plots do contain substantial independent information.

Because of the length of the videos, the RSA measurements are cleaner than they were for the images, and the differences between the groups parallel what we observed for the images. In the fear condition, socially liberal subjects do not change much on average, but conservatives' RSA falls, though the group-level difference is only borderline significant ($p = .053$). In the disgust condition, both groups show a strong increase in RSA, and there is some indication that conservatives may change more than liberals, though this difference is not statistically significant ($p = .16$).

As discussed in more detail in the SI, we replicated the RSA analysis using a more traditional peak-valley metric. The group-level differences in the fear condition were consistent across metrics, and conservatives showed somewhat higher RSA in the disgust condition, though the group difference was once again neither statistically significant or consistent between images and videos. Taken as a whole, the finding that conservatives had lower RSA in the fear condition is well supported and robust, while the finding that conservatives had higher RSA in the disgust condition is weaker and could be specific to the particular stimuli.

(a) Images



(b) Videos

Figure 2.4: RSA during the (a) image and (b) video stimuli relative to common baselines. The RSA baseline was calculated over the pre-stimulus questions in the first session or over the OS images in the second session.

2.3 Discussion

The results reported in OS were very different from the reactions of our subjects to those same images shown in the same order calculated by the same metric, but we found that the fixed ordering of the OS protocol coupled with the OS metric was likely the reason for this discrepancy.

We found that people with different levels of support for socially protective policies tended to react differently across the images, but the overall pattern is that more liberal subjects tended to react more strongly to all types of stimuli. We did not find a clear interaction between the type of stimulus and relative skin conductance reactions of liberal versus conservative subjects, though the RSA data suggests that the fear stimuli might differentiate between liberals and conservatives more than the others. This does not necessarily imply that fearful reactions are a more relevant driver of political ideology than disgusted or pleasant reactions since we are limited to a small sample of emotional stimuli, and the fearful stimuli that we use could simply generate more consistent reactions in people.

These results should not be interpreted causally, as there are simply too many mechanisms that could produce a positive correlation between social liberalism and reactivity. That said, there is clearly some connection between physiological reactions and political ideology, and it is the task of future research to parse out the precise mechanisms at play. Though in this paper we have employed the same measure of ideology as OS—support for socially protective policies—it is not clear whether this is more clearly connected to reactions than other measures of ideology. We have shown that there is some relationship between this ideological metric and reactions to images, but this certainly does not mean that other metrics are unrelated or even that social protection is the best metric to use. Our understanding of how biological factors relate to ideology is extremely limited, and an optimal ideological model for use in physiological or neurological studies has yet to be found.

Finally, external validity is an obvious concern here, especially since our sub-

jects' reactions to the OS protocol were so different from the OS subjects' under the same analysis techniques. It is possible that the factors that influence individuals' preferences vary across environments, so the relationship between social preferences and physiology may be modulated by other factors like locally dominant ideologies, group affiliations, race, ethnicity, or local events. Further research across a variety of populations, regions, or countries is needed to understand what facets of physiology and ideology are common across the population and which depend on other factors.

2.4 Materials

2.4.1 Subjects

40 subjects (17 female) participated in the experiment. The mean age was 41 (SD=13). About a third of the subjects were recruited through flyers posted at the Democratic and Republican field offices in Pasadena leading up to the 2010 elections. A small number of subjects were recruited through flyers posted at universities around Pasadena and Los Angeles, and then the rest were recruited through an advertisement on Craigslist. A preliminary filtering step was used to ensure that the subject pool was balanced between Democratic and Republican registrants.

Subjects were paid $30 per hour with a minimum of $50 if they took less than 1 hour 40 minutes, and parking was provided close to the laboratory. Subjects were scheduled for times that fit into their schedules, and they started the experiment between 10 AM and 5:30 PM. An effort was made to schedule subjects away from when they usually woke up or went to sleep.

Subjects were matched into groups by age, gender, and political party registration. Treatments were assigned randomly in each group, and for groups that had more than 3 subjects, the remaining treatments were randomized. In the second session, the treatments were rotated within each group to maintain balance across treatments, but the rotation direction was randomized between groups.

2.4.2 Stimulus materials

Treatment images came from the International Affective Picture System (IAPS) (Lang *et al.*, 2008), a set of standardized images commonly used in psychology research for eliciting various emotions. The disgust and fear conditions used sets of ten IAPS images each that were classified by (Mikels *et al.*, 2005) as either primarily disgusting or primarily frightening, and the control condition used a set of neutral positive images chosen from those classified by (Lang *et al.*, 2008) as moderate arousal and positive valence. The image codes are listed in the Supplemental information (Section 2.5).

The stimulus videos had also been used and characterized by others. For disgust, we used a clip that combined one minute from Trainspotting and one minute from Pink Flamingos;[2] for fear, we showed a clip from Silence of the Lambs; and for a control, we used a scene from Alaska's Wild Denali. The fear and control clips were edited according to the instructions in (Rottenberg *et al.*, 2007). The stimuli were all presented using the Psychophysics Toolbox Version 3 package for MATLAB (Brainard, 1997).

The OS image stimuli and presentation order were obtained from the OS authors.

2.4.3 Task and measures

Subjects came to the laboratory in two sessions; the first was in October or December 2010, and the second was in February and early March 2011. In the first session, they started with a variety of psychological batteries and a survey of their political preferences.[3] After those were complete, they were connected to a BioPac MP150 with modules to measure skin conductance on both palms, a 2-lead electrocardiogram, and respiratory effort. Subjects then answered 17 questions about their policy preferences and then viewed randomly assigned disgusting, frightening, or neutral positive images and videos. They then completed more policy preference

[2]This clip was kindly provided to us by Justin Feinstein from the University of Iowa.

[3]Specific details of these questionnaires are in the Supplementary Information.

questions as part of the priming effect experiment.

In the second session, subjects were connected to the physiological measurement equipment and were then shown the exact image set and ordering used by OS. They were then randomly rotated into one of the two treatment conditions that they did not see in the first session, and they viewed the new treatment's images and videos. As in the first session, subjects answered more policy preference questions, and then at the end, all subjects watched the neutral positive video to shift them towards a common emotional baseline and then viewed images from the two other categories, one of which they had seen in the first session.[4]

Skin conductance was measured using electrodes on the thenar and hypothenar eminences of the subject's non-dominant hand. Electrodes were also connected on the subject's dominant hand, but only the non-dominant signals are used here. A BioPac MP150 recorded both skin conductance channels, electrocardiogram, and respiratory effort signals.

There are many ways to estimate skin conductance response, but in this paper we examine only 3. The "OS" metric subtracts the log skin conductance from the 10 seconds before each image from the mean log skin conductance during each image. The other two metrics are calculated using the MATLAB package LedaLab (Benedek and Kaernbach, 2010a,b). The "phasic" metric extracts the phasic portion of the skin conductance signal from long-term trends and then measures the peak response during the stimulus away from the median signal in the 10 seconds before the stimulus. The "driver metric" calculates the SNS signal underlying the skin conductance responses by means of nonnegative deconvolution using a kernel function optimized to each individual's skin conductance characteristics. For the phasic and driver metrics, we masked out skin conductance measurements in the first second of each image or gap since SCR reactions to stimulus onsets are delayed by a few seconds, and so any spikes in the first second cannot be interpreted as reactions to the stimulus. The phasic and driver metric distributions are approx-

[4]The second session treatment appeared to impact subjects' reactions to these final images, so they are not used in the analyses.

imately lognormal with a few points at zero, so we transformed these signals by $\log(x + 10^{-3})$ in analyses. When multiple images are included in a measurement, the metrics are calculated for each image and are then averaged.

2.4.4 Analyses

SCR and processing was done in MATLAB 2010a, and RSA processing and statistical analyses were done in R 2.14.

Joint significance of group differences in reactions to all stimuli were calculated using permutation tests to account for the randomization structure. Each individual reacted to two measurements, so to account for interdependence between an individual's two measurements, we permuted support for socially protective policies and calculated the fraction of times that all the differences in group averages for the permuted data had the same sign and were as far in absolute value from zero as the actual data.

We convert the responses to the Wilson-Patterson questions into a binary variable rather than analyzing the sum of the responses as a continuous variable. This is necessary for the comparison with OS, and since it appears that most of the relevant variation in the Wilson-Patterson metric is around the median of the scale, we also treated the metric as continuous in our analyses. The reasoning for this is described in more detail in the Supplementary Information.

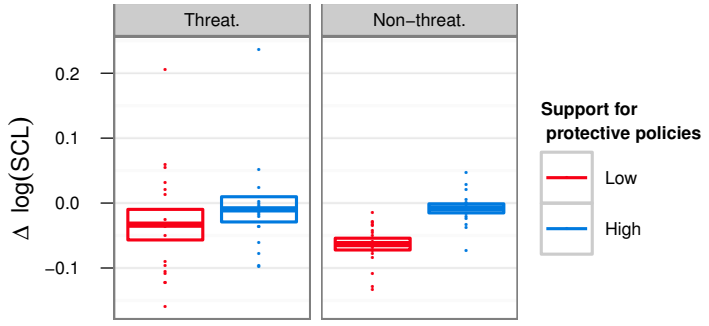
2.5 Supplemental information

2.5.1 IAPS image codes

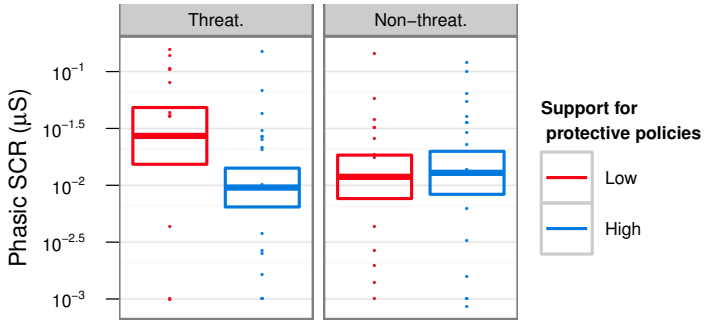
Table 2.2: IAPS codes for images used in the experiment

Disgust	:	7360 7380 9290 9300 9320 9330 9373 9390 9570 9830
Fear	:	6370 1110 1113 1301 1302 1930 1931 3280 5970 9600
Neutral Positive	:	7230 2650 7325 8470 2500 2370 1604 1810 1750 1500
True Neutral	:	2190 2440 2840 7000 7004 7006 7010 7020 7031 7035
		7050 7080 7110 7150 7175 7185 7187 7217 7491 7950

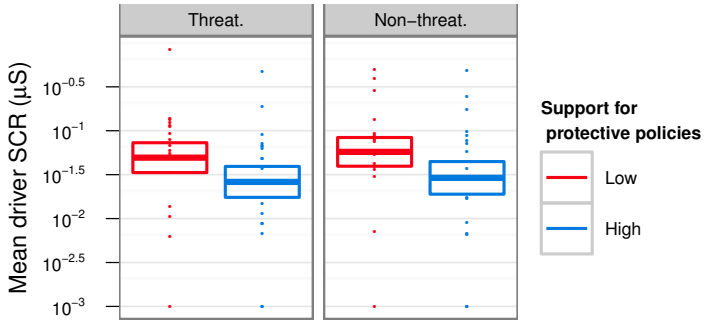
2.5.2 Comparison of SCR metrics



(a) Difference in logs

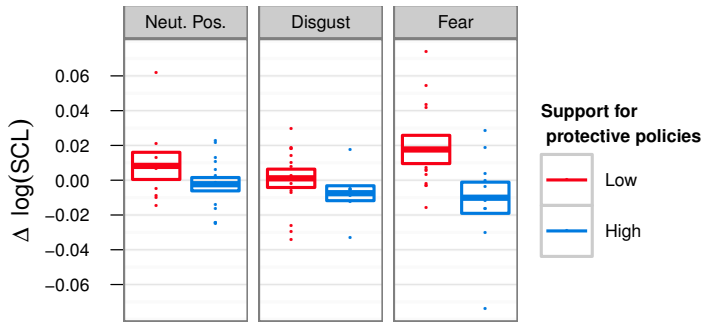


(b) Phasic baseline-to-peak



(c) Deconvolved driver signal

Figure 2.5: Comparison of our subjects' reactions to OS images under various metrics



(a) Difference in logs

(b) Phasic baseline-to-peak

(c) Deconvolved driver signal

Figure 2.6: Comparison of our subjects' reactions to IAPS images under various metrics

Note that the standard deviations of reactions to the IAPS images are smaller than for the OS images. There are two reasons for this. First, each data point under

the IAPS images is the average of 10 image reactions per individual while the OS images have only 3 per person. There are high levels of variability within an individual in how they react to each picture, and averaging over ten pictures rather than three improves the precision of individual estimates by a factor of about 1.8, and this then translates into tighter estimates of the group-level means. Second, the variability between images within an individual was lower on the IAPS images than on the OS images because the IAPS images were much more closely matched in terms of emotional specificity.

2.5.3 Baseline RSA

Baseline RSA was highly variable between individuals in both rounds, but there were no discernible group-level differences. Figures 2.7 and 2.8 show the group-level comparisons and also the correlation in baseline RSA across sessions.

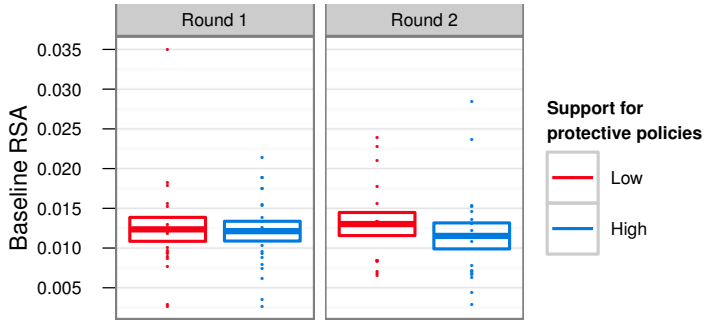


Figure 2.7: Comparison of the groups' baseline RSA calculated over the pre-stimulus questions in the first session or over the OS images in the second session

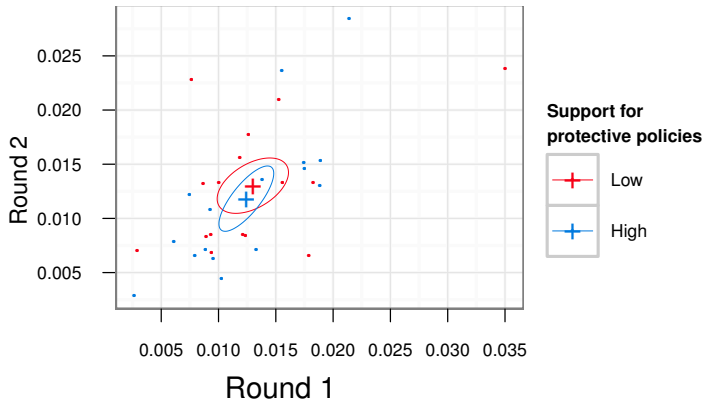


Figure 2.8: Scatterplot of each subjects' baseline RSA during the first and second sessions. + symbols give the group-level mean and are encircled by 95% confidence ellipses of the means based on a multivariate normal approximation.

2.5.4 Peak-valley RSA

Since the spectral RSA method is relatively new to the psychophysiology literature, we replicated some of the plots and analysis using an older peak-valley approach. To calculate this, we first located all peaks and troughs in the respiration data and calculated the instantaneous heart rate at each peak and trough. We generated a new stepwise time series with steps at each peak and trough and with levels equal to the difference between the instantaneous heart rate at the nearest peak and trough. We calculated the median of this time series over the stimulus display times after masking out the first 2 seconds that the stimulus was visible to account for response delays.

The peak-valley RSA values were similar to the spectral RSA values, though both are characterized by a moderate amount of noise. Fig. 2.10 compares the two during the baseline periods and Fig. 2.13 does the same during the videos.

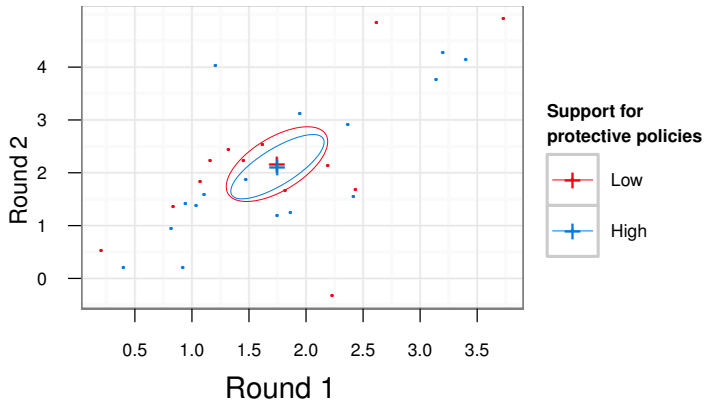


Figure 2.9: Scatterplot of each subjects' baseline peak-valley RSA during the first and second sessions

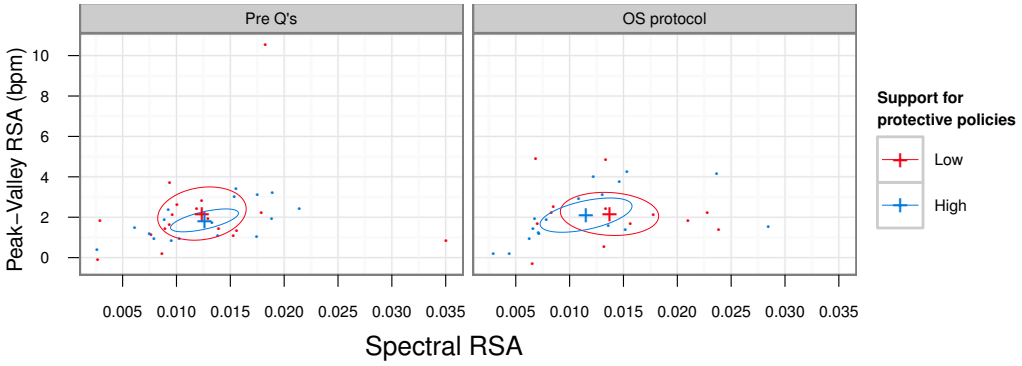


Figure 2.10: Comparison of spectral and peak-valley RSA during baseline periods

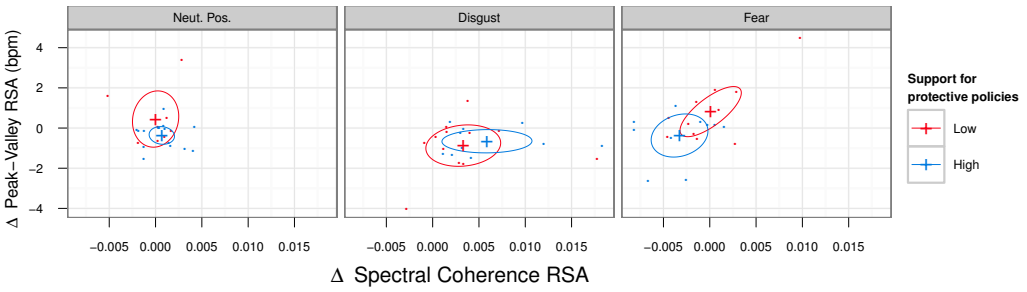


Figure 2.11: Comparison of spectral and peak-valley RSA during stimulus videos relative to common baseline

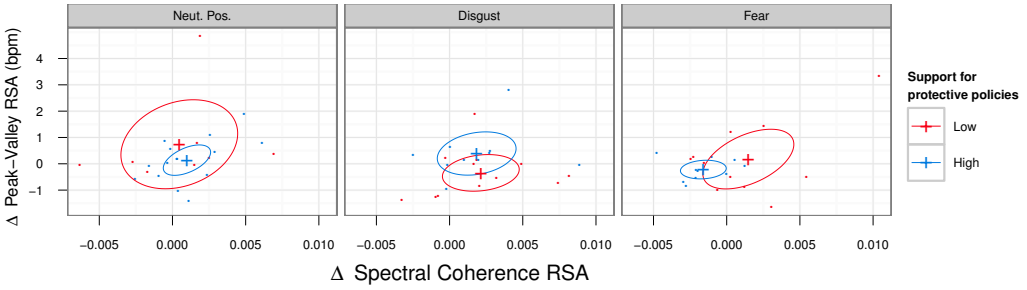
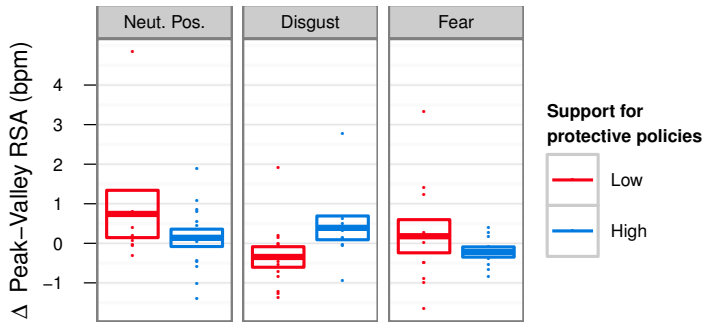
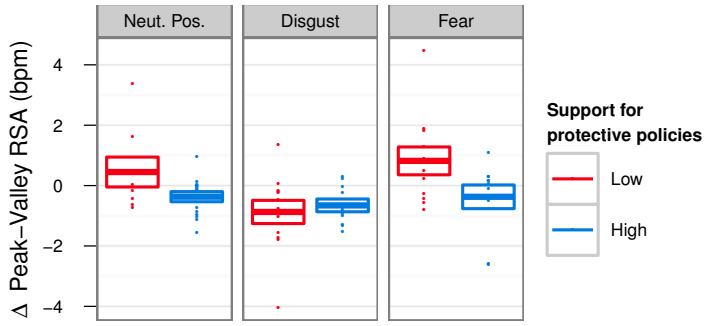


Figure 2.12: Comparison of spectral and peak-valley RSA during stimulus images relative to common baseline

Fig. 2.13 is the same as Fig. 2.4 in the main text but is calculated using peak-valley RSA rather than spectral RSA.



(a) Images



(b) Videos

Figure 2.13: Peak-valley RSA during the (a) image and (b) video stimuli relative to common baselines. The RSA baseline was calculated over the pre-stimulus questions in the first session or over the OS images in the second session.

The group-level differences for peak-valley RSA are borderline significant for fear videos ($p = .069$) and in the same direction but not significant for fear images. The difference for disgust images was also borderline significant ($p = .075$), but the disgust videos were not. This reverses the pattern of significance from the spectral RSA calculation, but the sign of the difference is the same across images, videos, and metrics, so the difference is suggestive but may be sensitive to particular stimuli.

Finally, the peak-valley RSA metric shows a significant difference in the neutral

positive videos ($p = .022$), and the sign is similar for images. These reverse the pattern under the spectral RSA metric where liberals had lower RSA (though the difference was not significant even when images and videos were grouped together). Because of the inconsistency between metrics and between stimulus types, and because the difference appears to be driven largely by two subjects in the tails, this is not clearly interpretable.

2.5.5 Wilson-Patterson discretization

We collapse subjects' responses to the Wilson-Patterson questions into a single binary variable that tells whether the subject gives responses that are more or less 'conservative' (as defined by OS) than the median. Normally, continuous variables should remain continuous for analysis, but that does not appear appropriate in this case. First, the data is not truly continuous in the first place; it is a vector of ordinal trinary variables (Disagree, Uncertain, Agree). There is no reason that the underlying latent ideology has a linear relationship with the sum of these trinary variables, nor is there a good reason that the relationship between reactions to emotional stimuli and latent ideology is linear, and so there is no particular reason to assume a linear relationship between the sum of the WP responses and reactions.

Of course, there is also no *a priori* reason that the relationship between two variables should be a step function, but the data appears to exhibit more of a step function than a linear trend. This is immediately clear from the scatterplot of the social WP index versus reactions to the images in Fig. 2.14. There are very few socially liberal subjects who have weak reactions to the stimuli, while among conservatives there are a mixture of strong and weak responses, or conversely, very few weak reactors who are socially liberal. OS found little difference between binary and continuous metrics for the WP index, but since that paper's sample was highly bimodal and had very few data points near the median of the WP scale, it had little statistical power to distinguish the two models.

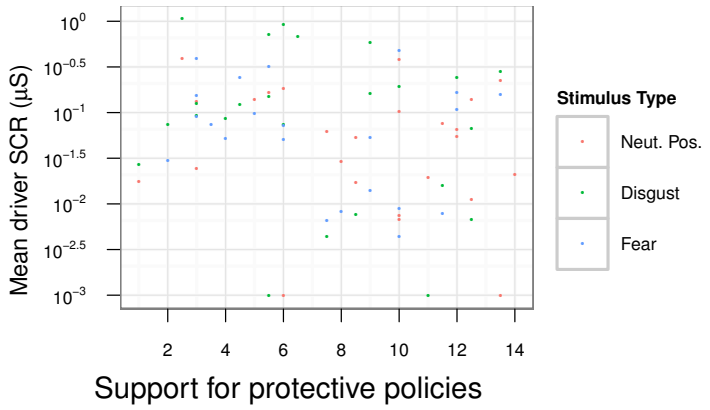


Figure 2.14: Scatterplot of driver skin conductance responses to videos versus additive Wilson-Patterson scale for socially protective policies

2.5.6 Replication details

Experiment details

All experiments were conducted in a single room at the California Institute of Technology. Initial surveys were done in a web browser, and the experimental stimuli were presented from MATLAB using the Psychophysics Toolbox Version 3 package (Brainard, 1997).

Questionnaire details

The pre-stimulus questionnaires included the PANAS, STAI, and a survey of political activities and ideology.

The political survey began with questions on voting history followed by the Wilson-Patterson ideological battery from (Oxley *et al.*, 2008). That was followed by the 30-question Moral Foundations Questionnaire and a 7-point branching party ID question.[5] We then asked quiz questions about political figures to gauge political knowledge, a series of ANES questions about participation during the 2008

[5]The 30-question Moral Foundations Questionnaire is available from `http://faculty.virginia.edu/haidtlab/mft/index.php`

election, and finally a series of demographic questions about age, education, income, race, marital status, children, and history of employment in the military, medical, or emergency response fields.

Chapter 3

Self-reported fear and item nonresponse on surveys

There are three classes of item nonresponse in surveys: structural nonresponse where the survey only asks some questions to a subset of respondents, order-dependent nonresponse like panel survey attrition and dropout, and respondent-driven where respondents are asked a question, but do not give an answer. Within the respondent-driven class, there are further differentiations based on the type of nonresponse; answering "Don't know" (DK) on a question is rather different from refusing to answer a question on privacy or propriety grounds. Our focus here is on DK responses; those where a subject is asked a question, they do not object to answering it, but they do not have strong enough thoughts, preferences, or information to give an answer.[1]

Survey researchers have known for a long time that DK responses are not completely random. Ferber (1966) analyzed a very large sample of consumer surveys and found that nonresponse rates varied with both respondent-specific and question-specific parameters. Francis and Busch (1975) followed this line of research for political surveys and found similar results, and concluded that simply ignoring non-substantive responses (DK, 'no opinion', etc.) or reassigning those responses to a fixed category both cause problems.

[1]This work was done in collaboration with Laura Loesch, R. Michael Alvarez, and Lonna Atkeson

More recently, Leigh and Martin (1987) showed that similar results hold for telephone surveys, and Rubin *et al.* (1995) demonstrated that censored data from election outcomes can be treated similarly to item nonresponse in surveys.

This line of research focused on understanding what observable respondent characteristics affected their propensity to skip questions or five non-substantive responses, but it did not address the possibility of latent biases in an individual's propensity to respond. More recently, Loosveldt *et al.* (2002) showed that item nonresponse is predictive of survey attrition in panel surveys after controlling for demographics, meaning there exist some set of respondent characteristics that predict item nonresponse and panel attrition that are not captured by standard demographics. This paper explores those latent influences.

3.0.7 Cognitive biases

Cognitive science has shown that a person is not as rational a decision maker as we would like to believe; heuristics—"rule of thumb" approximate solutions that allow for rapid processing in a complex world—bias our perceptions, mediating our interaction with the world (Gigerenzer and Goldstein, 1996).

The study of heuristic biases in voting and other political decisions is well situated within the framework of the Heuristic-Systematic Model of persuasion (Todorov *et al.*, 2002), which describes how superficially attended cues influence attitude change. The HSM specifically attempts to determine which circumstances elicit preferential use of either the systematic (based on careful attending of all relevant components of the persuasion message), or heuristic (based on only some informational cues, especially superficial cues) mode of message processing. While one mode may dominate processing, this dual-process model does not assume that only one mode is used at a time. When cues from both modes are congruent, simultaneous systematic and heuristic cues combine additively, further promoting the decision. On the other hand, the two systems can work against each other when systematic and heuristic cues are incongruent. In the incongruent case, more deliberately processed systematic cues will attenuate heuristic cues. However, if

those systematic cues are ambiguous about their persuasion message, heuristic cues will take precedence, even if this results in an individual being persuaded of something they would not agree with if the systematic cues had been more straightforward. In other words, when the facts are confusing, we rely more on our gut reaction to figure things out. Certain circumstances lead to the use of a certain mode of processing. Under conditions of reduced cognitive capacity, heuristic processing dominates; under conditions of increased motivation hopefully, voting is an example of this systematic processing dominates. Increased motivation is a function of the sufficiency principle, which states that processing effort (degree to which systematic processing is used) increases with an increased difference between someone's actual and desired confidence in a persuasive message. In short, if something is important to someone, they will desire high confidence in a persuasive message. Systematic processing of the persuasion message is more likely to yield a higher actual confidence in the message, which will match the high desired confidence. Although the HSM was designed to model persuasion, it has become a more general model of social information processing.

While exploring ways to directly test the above assumptions of the HSF, (Todorov *et al.*, 2002) developed the bias hypothesis of the HSM, which is specially situated to real-life application since it focuses upon how implicit, affectively driven processes bias persuasion. In this model, heuristic processing plays a key role. And real life—where we often receive mixed messages—is just the sort of ambiguous situation in which the HSM predicts that heuristic processing will come to the forefront. The hypothesis proposes that this heuristic bias in the face of ambiguity is independent of motivation. Hence, one may make a decision or be persuaded of something with which they might not consciously agree; this observation renders the sufficiency principle of the HSM powerless. The implications for survey research are significant: respondents may be so highly motivated to answer questions—or conversely, to skip questions—that they do so whether or not they actually have an answer.

3.0.8 Outline

This paper will introduce a set of questions about fear from Caltech's module on the 2008 CCES surveys, and will compare responses to those questions with "Don't know" responses on other survey questions. That comparison is split into three parts. We begin with simple bivariate comparisons and uncontrolled regressions, then we move to a traditional index-based regression approach to control for omitted variable biases and dig into demographic variation. We conclude with a newer modeling approach that uses latent variables and item response theory models to avoid some of the problems with index-based regressions. The paper finishes with a discussion of modeling techniques and the utility of using multiple analysis methods in explanatory research.

3.1 Measurement of fear

All the data in this paper comes from the 2008 CCES surveys, including both the Common Content and Caltech's module. The survey was administered online to a balanced sample of the US population, and it had a panel structure with two waves. The first wave was fielded before the 2008 general election, and then the second was given to the same respondents after the election. Just under 80% of the individuals from the first wave continued for the second.

For measuring respondents' fear, the following questions appeared as a matrix of radio buttons in the pre- and post-election 2008 CCES Surveys:

> **Question Text:** Listed below are some objects and situations. Using the scoring system below, rate each on the intensity of fear that you associate with that specific object or event.
>
> **Columns:** No fear, very little fear, A little fear, Some fear, Much fear, Great fear, Terror
>
> **Rows:** Spiders, Not being a success, Being a passenger in a plane, Suffocating

Rows were randomly rotated, and columns were randomly reversed. Figure 3.1 shows the distributions of responses to the four fear questions, split by responses on the pre-election and post-election surveys. Respondents covered the full range of possible responses, though as expected, more people report low levels than report high levels of fear.

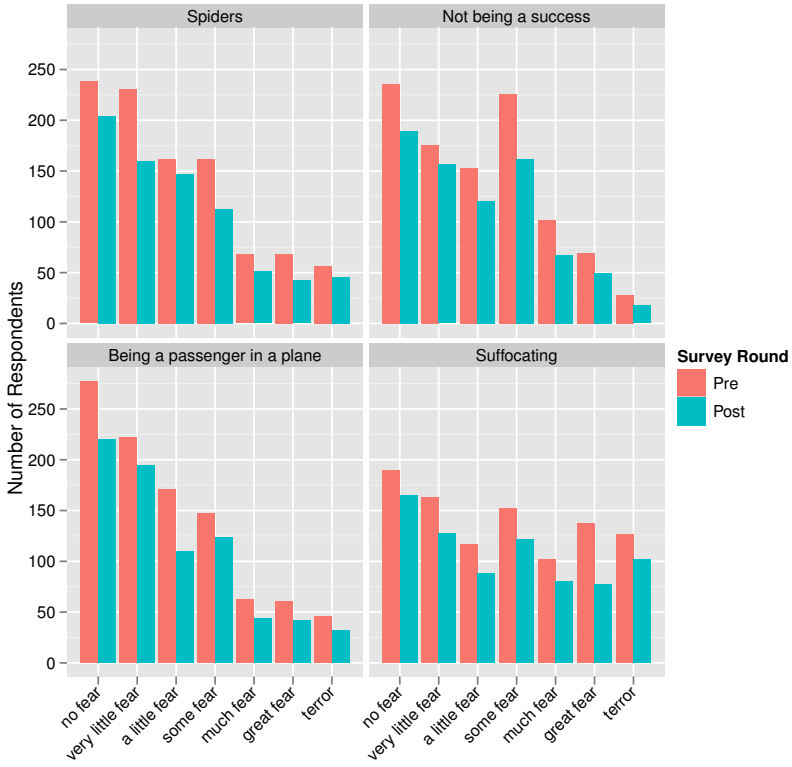


Figure 3.1: Distributions of answers to each of the 4 fear questions, split by responses on the pre-election and post-election surveys

Figure 3.2 shows the joint distribution of each variable's responses between the pre- and post-election surveys, and the correlations between the pre- and post-questions are shown in Table 3.1. Answers to each of the four questions were reasonably correlated between the two survey periods, with correlations ranging from a low of .61 for fear of not being a success to .77 for fear of spiders. The correlation with all fear answers grouped together was .70.

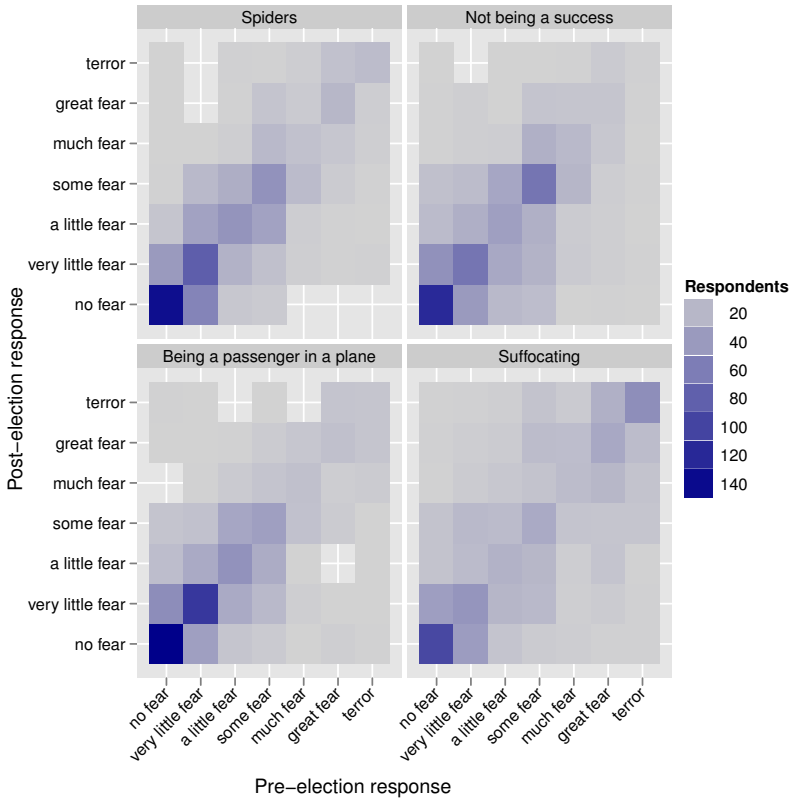


Figure 3.2: Joint distributions of the pre- and post-election answers to each of the 4 fear questions

Table 3.1: Correlation in answers to pre-election and post-election surveys by variable

Spiders	0.769
Not being a Success	0.605
Being a passenger on a plane	0.714
Suffocating	0.675
All Combined	0.704

The question therefore appear valid as a measurement tool, in that there is variation across the population, and there is consistency over time within individual respondents.

The 8 fear variables, four from each survey round, fall on an intuitive set of principal components. The first component puts positive weights on all the fear variables, and the second component differentiates between respondents who were more scared of not being a success than of being a passenger on a plane. The full set of factor loadings and component standard deviations are shown in Table 3.2, but for the purposes of this paper, we'll focus on only the first principal component.

Finally, though the questions are internally consistent, they have serious validity problems. In testing not reported in this paper, our measure of fear was not as predictive of political behavior as any of the biological or psychological experiments discussed in the introduction. Nor did it have as strong of an effect as that found in another fear-based NES study that by Skaperdas and Grofman (1995).

There are two primary problems with the questions we asked to derive the fear score. Firstly, they may have been too outright about asking about fear and therefore generated a response bias. For example, individuals who are socialized not to admit to various fears may have downgraded their actual fear in their report. Unfortunately, response bias is common in self-report surveys (Donaldson and Grant-Vallone, 2002) as individuals wish to portray themselves in a good light regardless of the confidentiality of the survey. One of the best ways to overcome this is to ask questions that are as non-indicative of socially loaded traits as possible. A second source that may have biased answers is how individuals read the individual questions—"How fearful are you of suffocating?" may have effectively asked one respondent if they are always terrified of suffocation while it may have asked another respondent if they would be terrified if faced with a situation in which suffocation seemed likely.

However, the second problem is that even if we can trust that respondents answered our question in a consistent manner, our questions may not have been appropriate to the task we wished to achieve. Fear of spiders; fear of suffocation;

Table 3.2: Principal component loadings and standard deviations for responses to fear questions in pre- and post-election surveys

Component	1	2	3	4	5	6	7	8
Pre: Spiders	0.3814	0.0070	-0.5748	0.1675	0.2908	0.0889	-0.0334	0.6344
Pre: Success	0.2953	0.5432	0.2535	0.1697	0.5280	-0.4307	-0.1580	-0.1886
Pre: Plane	0.3464	-0.4479	0.3064	0.2923	0.2821	0.1376	0.6225	-0.1187
Pre: Suffocating	0.3827	-0.0826	0.0952	-0.5964	0.2733	0.5175	-0.3155	-0.1998
Post: Spiders	0.3851	0.0217	-0.5687	0.1546	-0.2919	-0.0715	0.0343	-0.6422
Post: Success	0.2747	0.5713	0.2672	0.2200	-0.4470	0.4807	0.1498	0.1580
Post: Plane	0.3506	-0.4130	0.3140	0.3296	-0.3014	-0.1662	-0.6052	0.1208
Post: Suffocating	0.3927	0.0081	0.0951	-0.5666	-0.3265	-0.5060	0.3115	0.2365
Std. Dev.	1.8478	1.1560	1.0472	0.9607	0.7098	0.5614	0.4872	0.4160

fear of airplanes; fear of failure—all but the last of these pinpoint phobias. And phobias do not tap into pervasive behavior. They are situation-specific, and, as such, cannot reliably be expected to contribute to general political ideology, unless, of course, there is a party that promotes the extermination of all spiders.

What may instead be important to political behavior is if a person is extremely frightened of the majority of these things, as possessing that many phobias would be unusual. Indeed, we might expect such a person to be "scared" of nearly everything. This person would display a personality trait for anxiety, which would be relevant to political behavior. Indeed, Huddy *et al.* (2005) demonstrated that it is relevant to consider anxiety as distinct to specific fear in relation to political behavior: they showed that the perception of threat and experience of anxiety affected attitudes towards antiterrorism policies differently. Simply, being specifically afraid of something motivates behavior differently from being anxious. It behooves political scientists to choose which behavior they are interested in; for the original purposes—explaining general ideology, measuring anxiety seems most relevant. While high levels of fear as derived in this paper may point to anxious individuals, these questions are not actually very good.

The very things that make our fear questions problematic for actually measuring fear or anxiety–individual interpretations and social biases–can be helpful for tapping into where else those biases matter in survey research.

3.1.1 Predictors of fear variables

We modeled the first principal component of the fear questions based on available demographic covariates: race, income, education, age, gender, church attendance, and whether the respondent is a born-again Christian. We also checked whether ideology or party identification were related to fear reports. The results of these models are shown in Table 3.3.

In general, younger, poorer, and less educated respondents said they were more afraid. People who attended church said they had less fear, but Born-again Christians tended to be more fearful.

Ideology and party identification were not statistically significant, and their effects could be constrained to be relatively small. When a quadratic term was added for partisanship, it was not significant either. Partisanship and ideology also had little impact on the estimated effects of the other variables, so we can be relatively confident that the fear variables are not directly connected to ideology or partisanship.

Table 3.3: Fitted models of first principal component of fear questions

Dependent Variable	First Principal Component		First Principal Component		First Principal Component	
(Intercept)	10.752***	(0.571)	11.184***	(0.624)	11.245***	(0.662)
Ethnicity/Race: Black	0.305	(0.340)	0.176	(0.374)	0.159	(0.374)
Ethnicity/Race: Hispanic	0.069	(0.356)	0.030	(0.357)	0.054	(0.356)
Ethnicity/Race: Other	-0.449	(0.390)	-0.510	(0.395)	-0.494	(0.395)
Income Bin	-0.105***	(0.029)	-0.104***	(0.029)	-0.103***	(0.029)
Education	-0.243**	(0.081)	-0.263**	(0.082)	-0.250**	(0.081)
Age	-0.031***	(0.007)	-0.032***	(0.007)	-0.031***	(0.007)
Female	1.773***	(0.209)	1.776***	(0.209)	1.751***	(0.211)
Church: Never	0.612*	(0.281)	0.594*	(0.280)	0.614*	(0.280)
Church: Weekly	-0.404	(0.271)	-0.360	(0.273)	-0.380	(0.272)
Born-again	0.742*	(0.280)	0.777**	(0.285)	0.758**	(0.284)
Ideology	.	.	-0.058	(0.084)	.	.
Party ID	.	.	-0.036	(0.052)	-0.224	(0.199)
Party ID^2	.	.	.	.	0.021	(0.024)

3.2 "Don't know" responses

Across the pre- and post-election surveys, the CCES data has 139 questions with DK options[2], and of these, 128 were asked to at least 200 individuals could be included in our analysis. Since the 2008 CCES was administered online, the DK option was visible to respondents, so they all had an equal opportunity to skip those questions.

As would be expected, the questions had a wide range of DK frequencies. Questions with low rates of DK responses included one about the state of the economy, a post-election question about how people voted (conditional on them having voted), and a post-election question on how the respondent voted in the House race.

Questions with high DK frequencies included those about candidate placements on ideological scales, ones that ask about preferences on specific policies, a question about the previous day's Dow index, and one that asked respondents to estimate the unemployment rate. The distribution of DK frequencies across questions is shown in Figure 3.3.

3.2.1 Fear reports and "don't know"

The raw data hint at a connection between answers to the fear questions and DK responses. If we look at logit regression fits of 'don't know" responses to average fear response for each question separately, as in Figure 3.4, most of the curves have moderate upward slopes. The standard errors (not shown) of the individual curves are very large, but despite that, the average across all of the questions can be estimated very precisely.

Before we consider any potential confounding variables, there appears to be a statistically significant and relatively large connection between fear ratings and the frequency of DK responses.

However, we showed earlier that there was a strong relationship between fear

[2]We also include other phrasings of 'don't know' like 'not sure', 'I'm not sure', 'I do not know'.

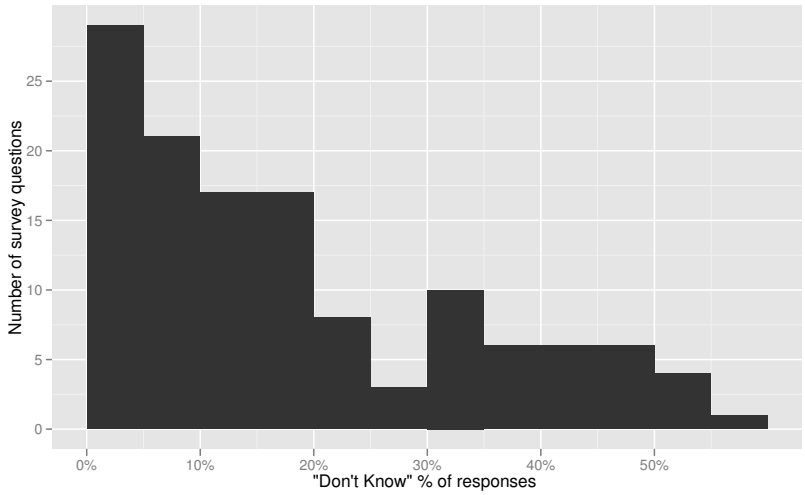


Figure 3.3: Distributions of DK responses across all questions that had DK or variants as possible answers

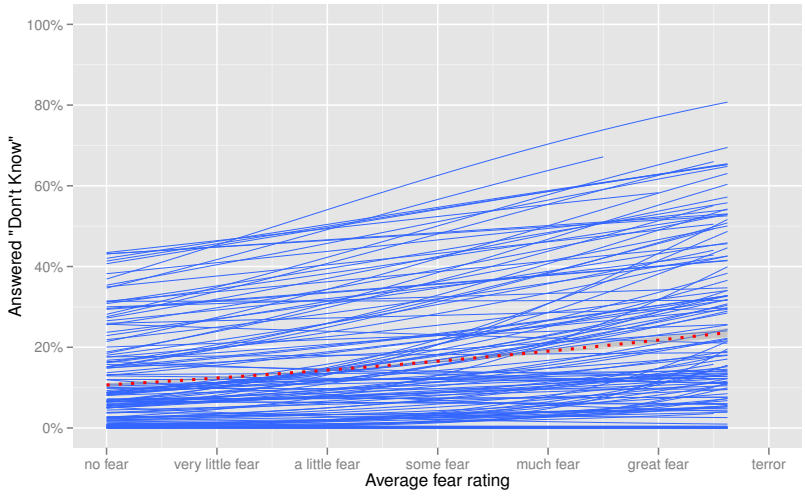


Figure 3.4: Fitted curves from logistic regressions of DK responses on an individual's average answer to the fear questions. Dotted red line is regression aggregated across all questions. The 95% confidence interval is shown in gray for the aggregated regression, but it is very narrow and may not be visible.

responses and a variety of demographic variables like age, gender, church attendance, education, and income. We would expect that some, if not all of these variables would also have an impact on an individual's ability or desire to answer a question, so there is a strong chance that the uncontrolled regressions suffer from omitted variable bias.

One interpretation of these results, which we will attempt to test later in the paper, is that individuals who are willing to admit being more afraid of things—or who do not feel social pressures to *not* admit fear—are also more willing to admit that they do not know things. In the US, social pressure to not admit fear is generally stronger for males than for females, and so we would expect that males who are unwilling to admit fear might do so because of social pressures, and those social pressures might also dissuade them from admitting that they do not know an answer.

The social undesirability of DK responses only matters if an individual does not actually have a clear answer to a question; social pressures do not matter if you would never answer DK regardless. Thus we would also expect that any of these social desirability differences would be modulated by an individual's political knowledge or interest.

Because of the danger of confounding variables, and our interest in interactions, the next two sections examine several other models that control for demographic effects and differences in political knowledge. The first section uses a traditional approach to model an index of DK responses based on demographics, a knowledge index, and a fear index. The second section uses a latent variable approach based on item response theory to avoid problems caused by simple additive indices.

3.2.2 Traditional model

The traditional approach to modeling this sort of data starts with indices of DK response frequency and of fear responses. The simplest option for the DK index would be to just add up all of an individual's DK answers.

However, in our data, respondents were asked varying numbers of questions.

Attrition between the pre-election and post-election waves of the CCES survey, randomized assignment, and branching assignment of questions all produce non-uniform missingness. Thus, instead of sums of DK responses, we have to work with averages. The missingness also means that the DK responses are not entirely comparable across individuals. For example, questions that were only asked to a few respondents tended to show very slightly higher rates of DK responses, so this averaging method is somewhat biased by construction. This will be remedied by the latent variable models that we introduce later.

The fear index was simpler. We simply averaged the 7-point response scales across all the questions an individual was asked. The fear questions were only missing when there was attrition between the pre and post waves, and since the same questions were asked in the pre and post waves, the fear scale does not suffer the same potential biases as the DK scale.

Finally, since an individual's interest in politics is likely to impact their desire to answer political questions, we also constructed a political activity scale. This added together indicator variables for whether a respondent had ever attended a campaign meeting, posted a political sign, persuaded someone else how to vote, donated money to a campaign, or worked or volunteered for a campaign. To aid interpretability, the fear and activity indices were rescaled to have mean 0 and variance 1 across the sample.

We then regressed the DK index on the fear index, the individual interest index, and demographic variables. Table 3.4 displays the fitted models with different indices included, and additional models interacting the indices with each other and with gender.

Not surprisingly, politically active individuals tend give answers to more questions. One standard deviation of the activity index translates into about 2% fewer DK responses (over the full activity scale, the effect is about 7%). When we look at the fear index on its own, the effect is just outside 5% significance, and the effect size about 3% over the scale's entire range (.6% per standard deviation). Considering that the average rate of DK responses across all questions is only

Table 3.4: Fitted models of DK index regressed on demographics and other indices. P-levels are denoted by *=.05, **=.01, ***=.001.

Variable	Base		Activity		Fear		Activity+Fear	
(Intercept)	0.221***	(0.015)	0.211***	(0.017)	0.217***	(0.015)	0.208***	(0.016)
Female	0.050***	(0.006)	0.046***	(0.006)	0.046***	(0.006)	0.041***	(0.007)
Education	-0.011***	(0.002)	-0.007**	(0.002)	-0.010***	(0.002)	-0.006**	(0.002)
Age	-0.343***	(0.088)	-0.271**	(0.094)	-0.342***	(0.090)	-0.224*	(0.095)
Age^2	0.048	(0.088)	-0.042	(0.094)	0.053	(0.088)	-0.028	(0.094)
Ideology	-0.008**	(0.003)	-0.011**	(0.003)	-0.008*	(0.003)	-0.011**	(0.003)
Black	0.017	(0.010)	0.023*	(0.011)	0.014	(0.010)	0.020	(0.011)
Hispanic	0.017	(0.012)	0.013	(0.013)	0.017	(0.012)	0.013	(0.013)
Other nonwhite	0.004	(0.014)	0.016	(0.016)	0.003	(0.014)	0.015	(0.016)
Income	-0.004***	(0.001)	-0.004***	(0.001)	-0.004***	(0.001)	-0.004***	(0.001)
No Church	0.009	(0.009)	0.009	(0.009)	0.009	(0.009)	0.009	(0.009)
Church Weekly	-0.010	(0.007)	-0.009	(0.008)	-0.009	(0.007)	-0.009	(0.008)
Born-Again	0.012	(0.007)	0.016*	(0.008)	0.011	(0.007)	0.015	(0.008)
Activity	.	.	-0.020***	(0.003)	.	.	-0.020***	(0.003)
Fear	.	.	.	.	0.006	(0.003)	0.009**	(0.003)
Activity:Female	.	.	.	.	.	.	.	.
Fear:Female	.	.	.	.	.	.	.	.
Activity:Fear	.	.	.	.	.	.	.	.
Activity:Fear:Female	.	.	.	.	.	.	.	.

Variable	Activity∗Gender		Fear∗Gender		Activity∗Fear		Activity∗Fear∗Gender	
(Intercept)	0.210***	(0.016)	0.218***	(0.015)	0.206***	(0.016)	0.205***	(0.016)
Female	0.046***	(0.006)	0.046***	(0.006)	0.040***	(0.007)	0.040***	(0.007)
Education	-0.007**	(0.002)	-0.010***	(0.002)	-0.006*	(0.002)	-0.006*	(0.002)
Age	-0.260**	(0.094)	-0.342***	(0.090)	-0.222*	(0.095)	-0.214*	(0.094)
Age^2	-0.038	(0.094)	0.052	(0.088)	-0.017	(0.093)	-0.008	(0.093)
Ideology	-0.011***	(0.003)	-0.008*	(0.003)	-0.011**	(0.003)	-0.011***	(0.003)
Black	0.022*	(0.011)	0.014	(0.010)	0.020	(0.011)	0.019	(0.011)
Hispanic	0.014	(0.013)	0.017	(0.012)	0.013	(0.012)	0.013	(0.012)
Other nonwhite	0.015	(0.016)	0.003	(0.014)	0.017	(0.016)	0.017	(0.016)
Income	-0.004***	(0.001)	-0.004***	(0.001)	-0.004***	(0.001)	-0.003**	(0.001)
No Church	0.010	(0.009)	0.009	(0.009)	0.008	(0.009)	0.008	(0.009)
Church Weekly	-0.009	(0.008)	-0.009	(0.007)	-0.009	(0.008)	-0.009	(0.008)
Born-Again	0.015	(0.008)	0.011	(0.007)	0.014	(0.008)	0.014	(0.008)
Activity	-0.014**	(0.004)	.	.	-0.021***	(0.003)	-0.015***	(0.004)
Fear	.	.	0.007	(0.005)	0.008*	(0.003)	0.010*	(0.005)
Activity:Female	-0.015*	(0.006)	.	.	.	.	-0.012	(0.007)
Fear:Female	.	.	-0.001	(0.006)	.	.	-0.006	(0.007)
Activity:Fear	.	.	.	.	-0.008*	(0.003)	-0.001	(0.005)
Activity:Fear:Female	.	.	.	.	.	.	-0.015*	(0.007)

14.8%, though, the fear index effect is potentially large.

Looking at the model that interacts Activity with gender, we see that the effect of political activity on answering questions is twice as high among females as among males (2.9% and 1.4% per standard deviation of Activity). The estimated effects are nearly unchanged when fear interactions are included, though the standard error estimates widen enough to just drop the gender interaction below 5% significance.

When we bring Fear into the picture, we first notice that when Activity is included in the model, the Fear index tends to reach significance. However, the point estimates do not change dramatically between models, so the variation in significance across models is not particularly informative. The main effect of Fear is relatively consistent across the models at about .8% per standard deviation.

Since social biases can only affect DK answers when the respondent does not actually have a clear answer to give, we would expect that Fear would matter more for people who are not particularly interested in politics. This is exactly what we see. In the Activity∗Fear model, people who are 1 standard deviation above the norm on Activity are not affected by Fear, while people who are 1 standard deviation below on Activity have a very large Fear effect–about 1.6% per standard deviation of Fear.

Finally, since we expect that males' Fear indices are influenced more by social pressures than those of females,[3] we expect that the interaction between fear and activity will be stronger among males. The regressions show just the opposite—the interaction between activity and fear that we just described appears to be limited to females.

Males have an effect from Fear (higher Fear produces more frequent DK answers), but they have almost no interaction effect between Activity and Fear. For females, the effect of activity is slightly greater than for males as in the previous

[3]Or at least, males are influenced towards over-reporting relative to females, meaning that the difference in social bias between males and females is positive. We have no way to differentiate this from the case where females are biased to over-report.

model (though the gender difference does not quite reach significance), but the effect of Fear is much greater. Females who are active in politics are influenced little by Fear, but those who are not active are strongly affected. Inactive females with high Fear answer "Don't know" more often, while those with lower Fear answer "Don't know" less often.

Among females, the data are very consistent with a social bias explanation: Females who are biased to admit fear are also biased to admit not knowing answers. Females who are active in politics (and who likely have clearer and stronger opinions) are not affected by the biases.

Inactive and active males have similar effects of fear, which points away from there being social biases being at play. If the connection between fear and DK was rooted in social biases, we would expect that fear would matter more among inactive males, but we do not see that in the data. We are left with an open question for why males have a connection between fear and DK responses. It is unlikely that the connection comes from social biases, but we have no clear signals about where else it might come from.

3.3 Latent variable models

The above analysis uses indices of fear answers, DK responses, and activity in simple OLS models. For such indices to actually represent an underlying latent variable, we have to assume a strong and facially implausible structure on the relationships between the component questions and the distributions of the latent variable across the population. A more flexible approach uses item response theory and treats each question in a set as a separate model (often logits for binary data), with separate intercept and slope terms for each outcome variable.

This relaxes the assumption that DK responses are the same across questions. Questions can be easier or harder for people to respond to, and DK's to a particular question can be informative about the latent variable of interest, or they can just be question-specific noise. Most importantly, we do not have to figure these things

out for ourselves–the data give enough information to fit the question-specific parameters.

The 2008 CCES survey also included a set of quiz questions about facets of an individual's political knowledge. The questions included things like which party was the majority in the US House and Senate and the party of their senators and governors. There missingness patterns of these questions make them unusable for the traditional model, but the latent variable model estimation procedure handles sparse data directly, so we can use these questions with the comparatively weak assumption that data is missing at random conditional on other variables included in the model.

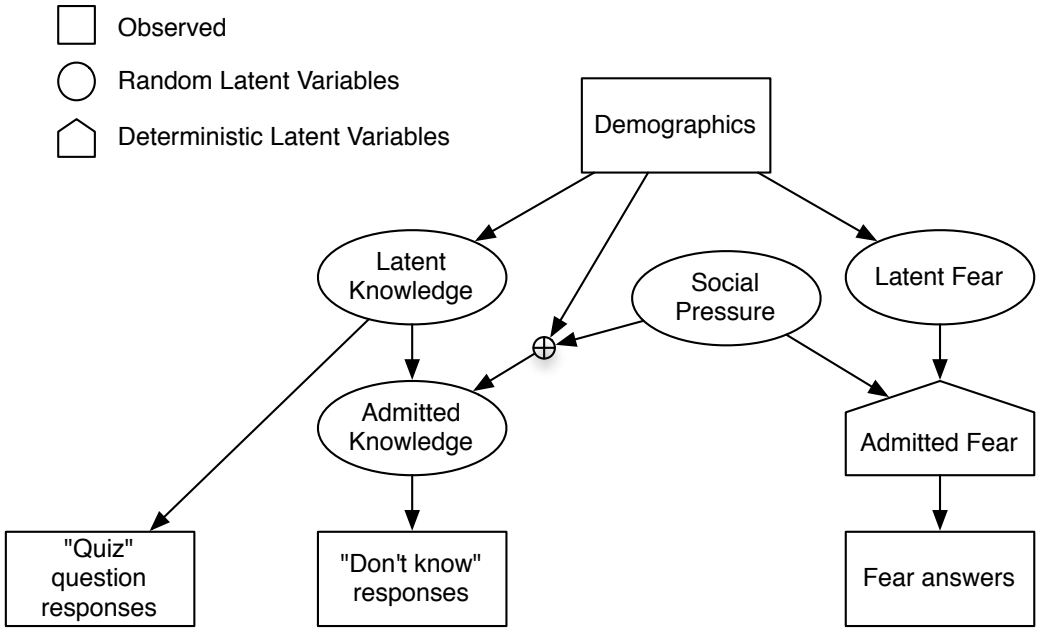


Figure 3.5: Graphical representation of the latent variable model. Rectangular nodes are observed data, oval nodes are unobserved latent variables with a random component, and pentagonal nodes are unobserved latent variables that are fully deterministic.

A graphical representation of our ideal model is shown in Figure 3.5. Latent fear represents the answers that people would give in the absence of social pres-

sure, and then admitted fear is just the latent fear biased by social pressures. Admitted fear is deterministic for model identification purposes, and—to simplify interpretation—'social pressure' can be interpreted as social biases in fear reporting.

Correct responses to quiz questions are assumed to be unbiased signals of latent knowledge, meaning, for example, that whatever social pressures might modulate a respondent's willingness to admit fear or DK will not affect whether they know which party has that House majority. Admitted knowledge is latent knowledge after social pressures have influenced how likely someone is to admit that they do not know something. We include an interaction with demographics, but only let the demographic variables modulate the strength of social pressure's influence on admitted knowledge. That restriction balances the desire to split apart the effects on different groups against a need to keep an already complicated model as simple as possible.

This model is too much of a departure from the traditional approach for us to directly compare the results in a meaningful way. There would be no way to pin down what facets of the model are responsible for whatever differences we might observe. Instead, the model will be built up in stages to highlight the impact of each bit of added complexity.

3.3.1 Separate models for fear and item nonresponse

We will begin by examining two separate unidimensional item response theory models. There are continuous latent variables called "Fear" and "INR" with information about an individual's level of fear or likelihood of answering DK on survey questions. We use cumulative link logistic models with varying discrimination, difficulty, and cutpoints on each fear question to connect responses to the continuous latent "Fear" variable. Item nonresponse is binary, so item nonresponse probabilities are modeled as binary logits with varying difficulty and discrimination between questions. Finally, both "Fear" and "INR" have linear demographic predictors and i.i.d. noise between individuals. The correlation between "Fear"

and "INR" is not explicitly modeled here.

For clarity, the responses for each individual i are modeled as shown below:

$$\text{Fear:} \qquad \text{response}^F_{qi} = 1 + \sum_{j \in 1,\ldots} \left\{ \gamma^F_q \text{Fear}_i + \epsilon^F_i > \text{CP}^F_{qj} \right\} \tag{3.1}$$

$$\epsilon^F_{qi} \overset{i.i.d.}{\sim} \text{Logistic}(0, 1) \tag{3.2}$$

$$\text{Fear}_i \overset{i.i.d.}{\sim} \text{Normal}(X_i \beta^F, \sigma^F) \tag{3.3}$$

$$\text{INR:} \qquad \text{response}^I_{qi} = 1 + \left\{ \gamma^I_q \text{INR}_i - \text{CP}^I_q + \epsilon^I_i > 0 \right\} \tag{3.4}$$

$$\epsilon^I_i \overset{i.i.d.}{\sim} \text{Logistic}(0, 1) \tag{3.5}$$

$$\text{INR}_i \overset{i.i.d.}{\sim} \text{Normal}(X_i \beta^I, \sigma^I) \tag{3.6}$$

CP^F_{qj} and $-\text{CP}^I_q$ are the fear question cutpoints and the INR logit intercept terms, γ denotes discrimination terms, X is a matrix of demographic variables where each column is centered and scaled to variance 1, and $\{\ldots\}$ is an indicator function. To simplify comparisons between models, the latent spaces are transformed after sampling so that "Fear" and "INR" have mean zero and variance 1. That rescaling is left out of the mathematical definition above for clarity. All the responses are coded as integers starting with 1 to simplify comparisons with the model fitting code. Because we have strong priors that the latent "INR" and "Fear" parameters should all have positive effects on each question, we constrain the γ terms to be positive.

This model was coded in the Stan modeling language (Hoffman and Gelman, 2011; Stan Development Team, 2013a,b) and run from R (R Core Team, 2012). Weakly informative double-exponential priors were used on for β^F and β^I. 4 chains were used for each model, and 500 draws were taken from each chain following an initial 100 adaptation steps.

There are two main features of substantive interest in this model. The first are the hierarchical predictors–how the various demographic terms impact fear questions and item nonresponse. The second is the amount of correlation between $(\text{Fear}_i - X_i \beta^F)$ and $(\text{INR}_i - X_i \beta^I)$ that was not explained by hierarchical predictors.

Taken together, these features give us a sense for how tightly item nonresponse is linked to self-reported fear, and a sense for whether the connection is being captured by available demographic variables.

Figure 3.6 shows the fitted parameter from both models. The most striking feature is the amount correlation in the parameter estimates between the two models. The fear and INR estimations were independent of each other, and so the connection between self-reported fear and item nonresponse is not just a feature of individuals—the same general group-level patterns hold across both measurements.

In terms of specific parameters, note that demographic variables are scaled to mean 0 and variance 1, and the latent space is rescaled to have variance 1 on all dimensions, so the proportions of variance explained by each variable are proportional to their square distance from 0. Females have much higher levels of both fear and propensity to skip questions, but the effect of gender still only accounts for approximately 10% of the variance in both fear and item nonresponse ($0.32^2 = 0.1024$). The next most important variables are: Education, which decreases INR substantially and has only a small impact on Fear; Age, which probably increases Fear more than it increases INR; and Income, which probably increases INR slightly more than it does Fear. We say 'probably' because the 40% ellipses are very close to or overlap the 45° line, and so the relative differences have substantial uncertainty.

The impacts of other variables are less clear. Conservative ideology tends to predict lower INR and slightly less fear, as does weekly church attendance (relative to the baseline occasional church attendance). African-Americans and Born-Again Christians had slightly higher fear and INR, but the effects are noisy relative to their sizes.

The model differentiates specific individuals reasonably well. Figure 3.7 shows the mean latent fear and INR estimates for each individual, plus a representative 95% joint credible ellipse for an individual at $(0, 0)$. There is still substantial noise in the individual-level estimates, so the group-level parameters are helpful, but there is still a lot of individual heterogeneity that the group-level parameters do

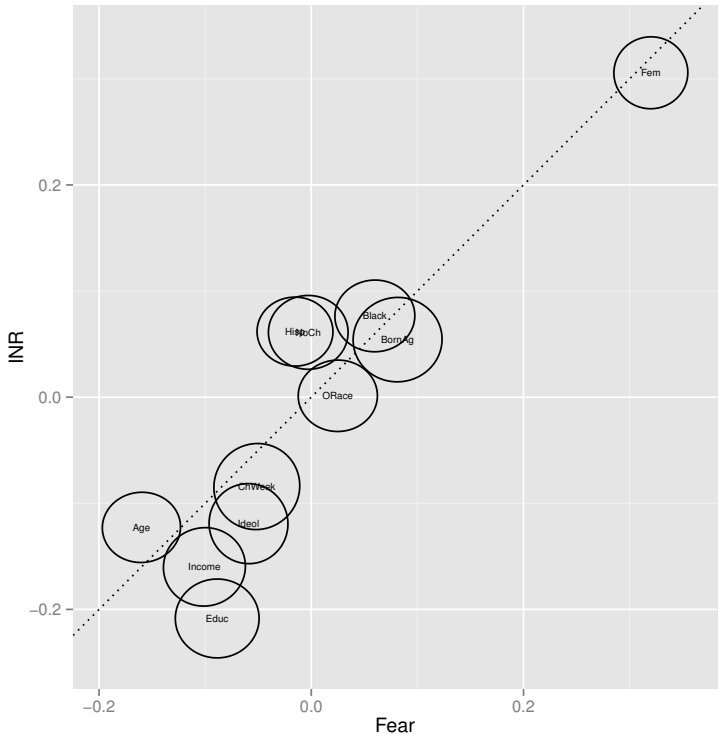


Figure 3.6: Fitted parameters of the separate fear and INR models. The overlapping labels are for Hisp and ORace. The circles are the joint 40% confidence ellipses for each parameter in the fear and INR models. 40% ellipses correspond to $t = 1$, and they are used to avoid excessive overlap in the visualization. 90% confidence ellipses are about 2.1 times larger, and 95% ellipses are about 2.4 times larger. The ellipses are based on the mean and covariance of each parameter's draws, which is an approximation of the actual joint distribution of the draws. The ellipses are circular because the models are fitted independently and do not incorporate correlated error. The dotted 45 ° line shows where the percentage of variance explained for Fear equals the percentage of variance explained for INR.

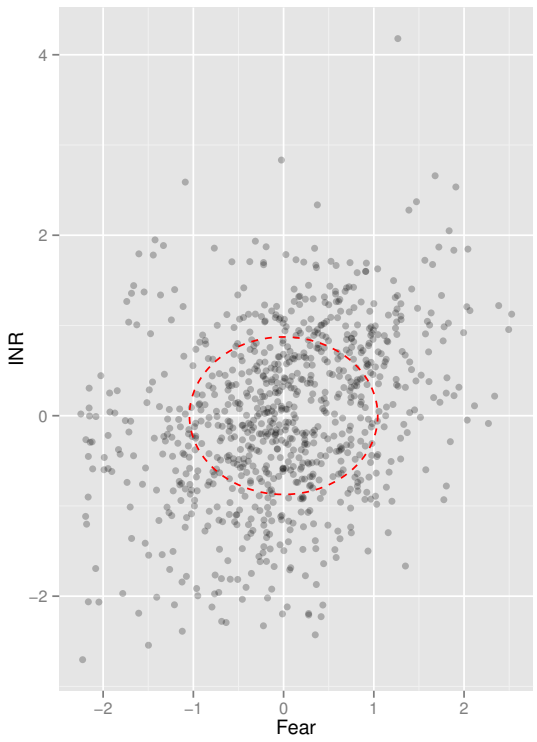


Figure 3.7: Mean latent fear and INR estimates for each individual. The dotted red ellipse is a representative 95% joint credible ellipse for an individual at $(0, 0)$.

not capture. The latent INR parameter is also easier to predict at an individual level than the fear parameter thanks to the large number of questions that people could skip, and so that also translates into more variation in the mean estimates on the INR axis. The draws at each stage were scaled so that the latent terms had variance 1, but variation between draws in each individual's estimate will tend to shrink each individual's mean towards 0—and at the extreme with complete noise, every individual's mean would be at $(0, 0)$.

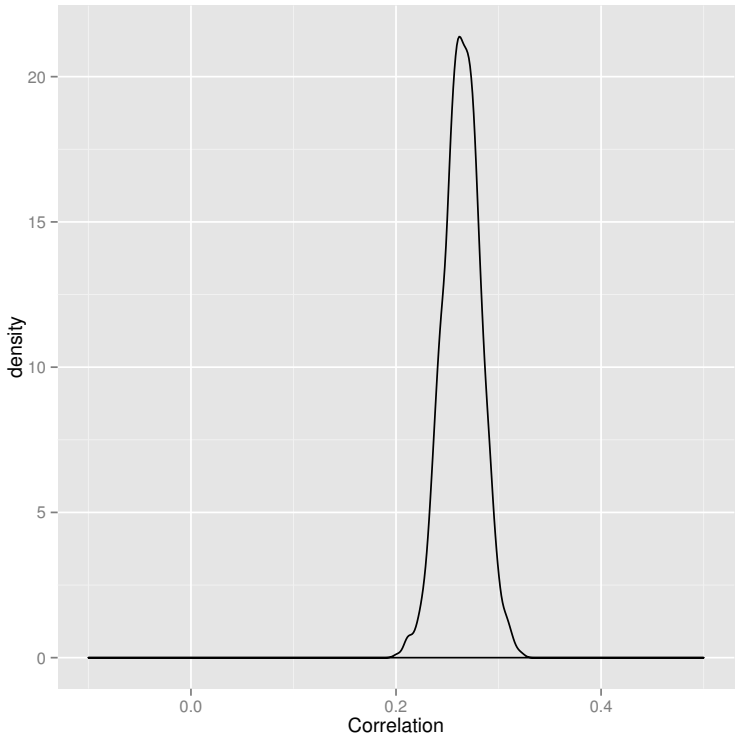


Figure 3.8: Distribution of correlations in latent fear and INR parameters across draws from posterior

The correlation between latent fear and INR parameters is apparent in the scatter plot of latent variable means (Figure 3.7), but Figure 3.8 shows the distribution of this correlation more clearly. The 95% credible interval for the correlation is $[.228, .300]$, but a substantial amount of this is accounted for by the demographic variables. For example, 'Female' accounts for about 10% of the variance in Fear

and INR, and the other variables tend to impact Fear and INR similarly, so the end result is a high level of correlation.

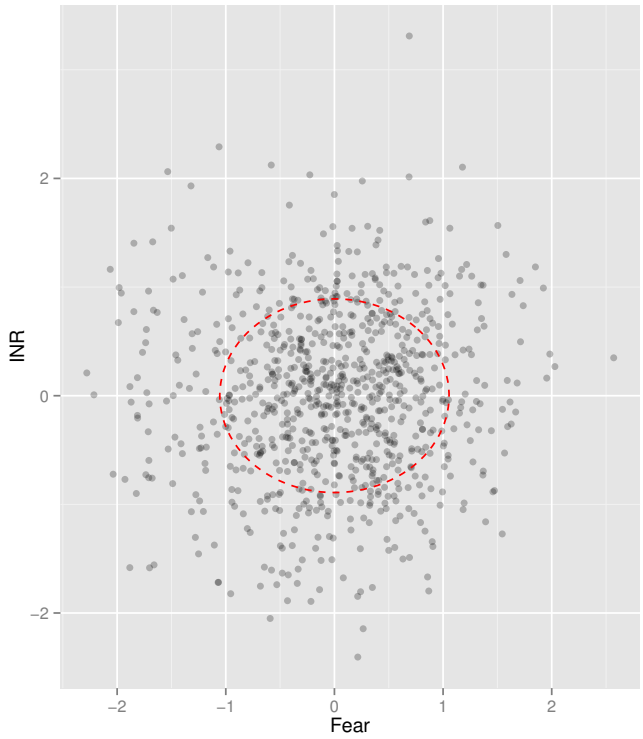


Figure 3.9: Mean individual-specific variation in latent fear and INR terms. The dotted red ellipse is a representative 95% joint credible ellipse for an individual at $(0, 0)$.

When we remove the demographic effects and look at just the individual-specific variation in Figure 3.9, the amount of correlation is obviously smaller, and is not immediately obvious in the scatterplot. But the distribution of correlations after removing demographic effects (Figure 3.10) shows that there is still a positive correlation between INR and Fear, so the connection is not explained entirely by the demographic variables included here. The 95% credible interval on the latent noise correlation is $[0.008, 0.091]$.

Note that since the models were estimated independently, the correlation in latent noise underestimates the actual correlation. The amount of the underesti-

mation depends on the amount of actual correlation (large correlations will tend to be underestimated more), but that will be addressed by the unified model in the next section.

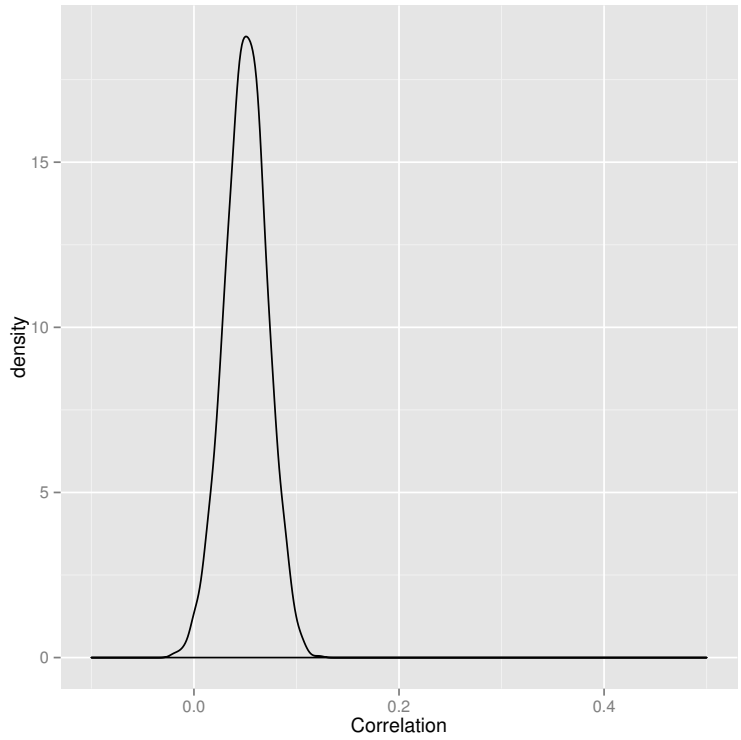


Figure 3.10: Distribution of correlations in individual-specific latent fear and INR variation across draws from posterior

Gender differences

For illustration of the effect sizes of demographic variables, Figures 3.11 and 3.12 show the same latent variables and individual variation as 3.7 and 3.9, but with individuals colored by gender.

The distributions of males and females are clearly different, and individuals towards the positive ends of the INR and Fear scales are almost all women, while the low end of both scales is occupied almost exclusively by men. There are not clear problematic patterns of heteroskedasticity between males and females nor

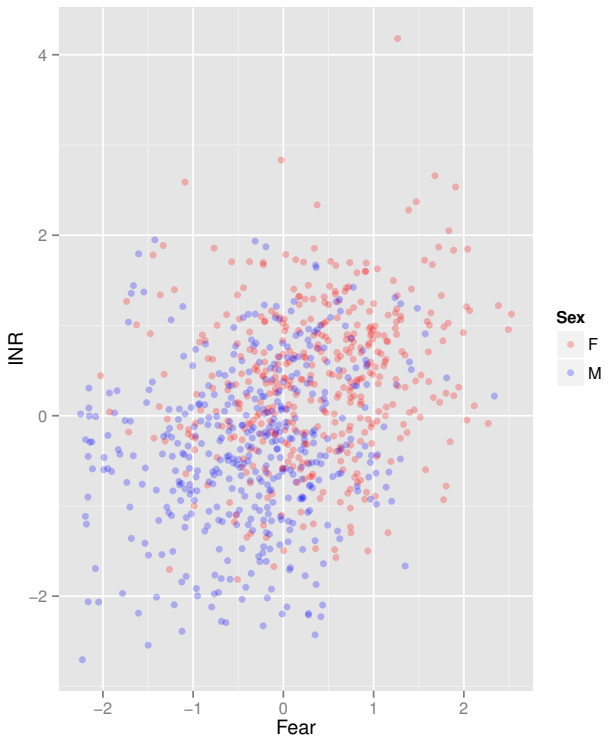


Figure 3.11: Mean latent fear and INR estimates for each individual, colored by gender. Red dots are females, blue ones are males.

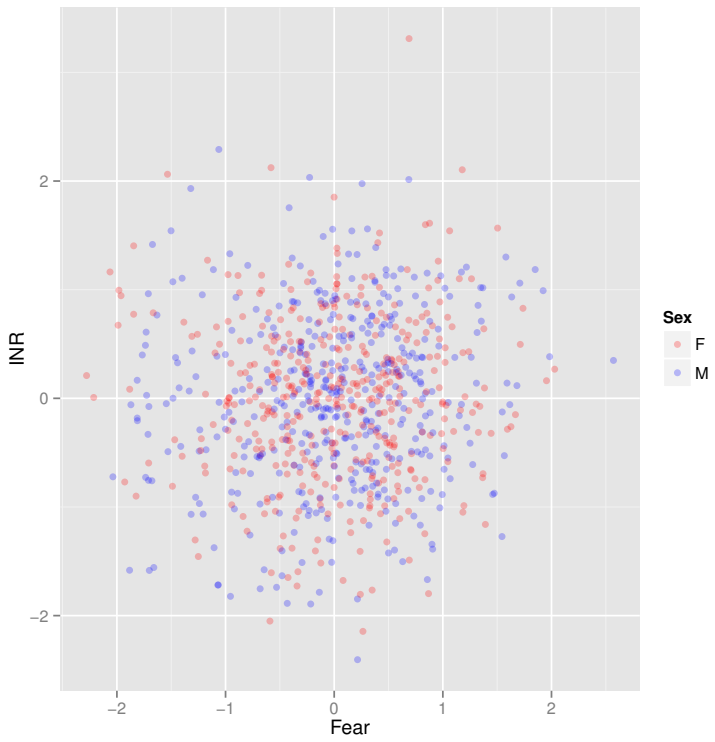


Figure 3.12: Mean individual-specific variation in latent fear and INR terms, colored by gender. Red dots are females, blue ones are males.

'chunky' distributions that would signal estimation issues or misspecification.

3.3.2 Unified model

There is obvious correlation in each individual's fitted ϵ terms in the previous model, so the next model iteration will capture that explicitly. The basic structure of the model is the same, but now the latent variables are drawn according to

$$(\text{Fear}_i, \text{INR}_i) \sim \text{MVN}(X\beta, \Omega) \tag{3.7}$$

Since the unexplained correlation was rather small between the independent models described above, we would not expect large changes in the demographic parameters when the correlation is modeled explicitly, and we would expect a small increase in the unexplained correlation.

The fitted demographic parameters for the unified model (Figure 3.13) are nearly identical to those from the separate model (Figure 3.6), which is to be expected since the noise correlation has only an indirect impact on those estimates. The main impact is that the correlated noise produces a slight correlation in the uncertainty of each demographic variable's INR and Fear parameters.

The latent variable correlations in Figure 3.14 are increased slightly (around 0.02) relative to the same plot for the separate models in Figure 3.8, and the 95% credible interval on the latent correlation moves up to $[0.242, 0.325]$ from $[.228, .300]$ for the separate model. The overall latent correlation results from similarities in demographic parameters plus then correlation in the latent noise, but the change in the latent noise correlation explains most of the shift between the separate and unified models. Figure 3.15 shows the distribution of this correlation, and the shift is again only slight, with the 95% credible interval moving to $[0.030, 0.125]$ from $[0.008, 0.091]$ in the separate models.

Overall, the correlation in the latent noise is small enough that adding it into the model does not make a substantively important difference, but it remains to be seen which steps later in the model's expansion will matter.

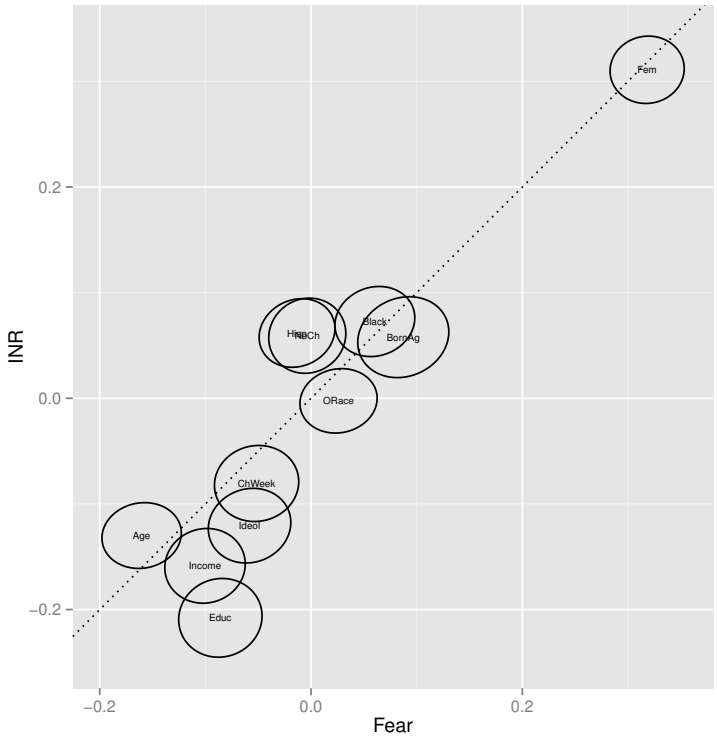


Figure 3.13: Fitted parameters of the unified fear and INR model. The circles are the joint 40% confidence ellipses for each parameter in the fear and INR models. 40% ellipses correspond to $t = 1$, and they are used to avoid excessive overlap in the visualization. 90% confidence ellipses are about 2.1 times larger, and 95% ellipses are about 2.4 times larger. The ellipses are based on the mean and covariance of each parameter's draws, which is an approximation of the actual joint distribution of the draws.

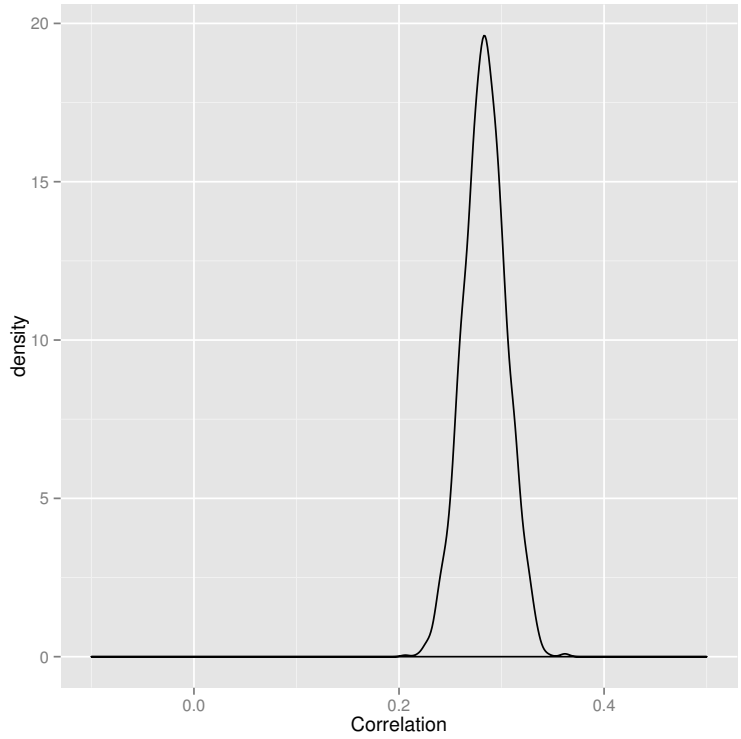


Figure 3.14: Distribution of correlations in latent fear and INR parameters across draws from posterior of the unified model

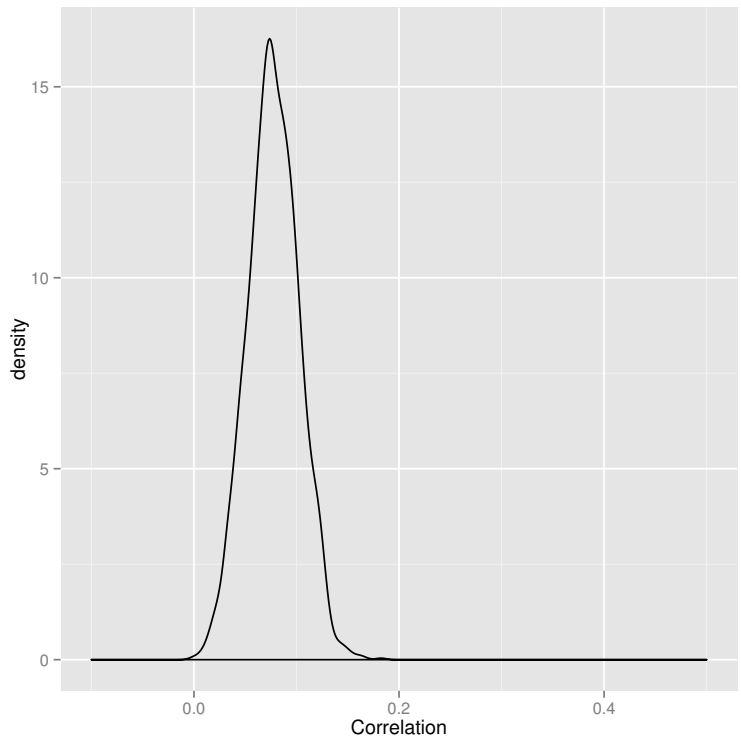


Figure 3.15: Distribution of correlations in individual-specific latent Fear and INR variation across draws from posterior of the unified model

3.3.3 Incorporating quiz questions

The next step in the model's expansion is to incorporate quiz questions that ask respondents to identify which parties are in the majority in the US House and Senate, the parties of their governor, senators, and representatives, and the race of their representative. These questions are an imperfect measure of political information, but they are the best objective measures that were available in the survey.

Methodologically, the only difference between this model and the unified one is the addition of a third latent dimension, Quiz, that is higher for people who answer quiz questions incorrectly. We reverse the 'natural' direction of a knowledge scale to simplify visual comparisons between the latent factors and their effects.[4] The primary purpose of this section is to compare the behaviors of the Fear, INR, and Quiz components of the model.

The same plots will be shown as before, but now we have 3 facets per plot to capture all the bivariate comparisons between the 3 variables. The demographic parameters shown in Figure 3.16.

The demographic predictors of Fear and INR in Figure 3.16 have a relationship very much like the unified model just discussed, and the Quiz predictors show a remarkably similar pattern. All three of latent predictors have quite similar patterns of demographic parameters, in that a positive parameter on one dimension is likely to be positive on all the others, and similarly for negative parameters. One surprising feature is that the Quiz and Fear parameters are more similar to one another than Quiz and INR. Quiz is an objective measure—high values are people who get factual questions wrong– whereas INR and Fear reflect choices made by respondents. We would expect biases in the choice process to impact Fear and INR, but not Quiz, and such shared biases would seem to lead to more covariance in the parameter estimates. The difference here is a bit too noisy to analyze deeply,

[4]Correct responses to the quiz questions are highly negatively correlated with both Fear and INR, so flipping the sign makes all the correlations positive to simplify magnitude comparisons, and it makes the graphs of parameter estimates line up in a much more readable way.

but it highlights the possibility that Fear may impact INR partially by influencing information collection about politics.

The Quiz latent variables are more difficult to predict at the individual level than Fear or INR because there are only 7 binary outcomes on the Quiz dimension. This is visible in Figure 3.17 by the narrow distribution of individual means on the Quiz dimension and the width of the red representative uncertainty ellipse on the Quiz dimension—there is roughly twice as much uncertainty in the Quiz estimates as in the INR estimates.

Though the parameters for Quiz and Fear were very similar, there is a huge amount of correlation in unexplained individual variation on Quiz and INR (Figure 3.20 and Table 3.6), and this dominates the parameter-based correlation between Fear and Quiz when we look at overall correlations in the latent variables (Figure 3.18 and Table 3.5). It appears that there are strong common factors connecting Quiz and INR beyond the demographic variables included here.

Table 3.5: 95% credible interval bounds of correlations in latent Fear, INR, and Quiz parameters across draws from posterior

Variables	2.5% Quantile	97.5% Quantile
INR, Fear	0.242	0.325
Quiz, Fear	0.257	0.422
Quiz, INR	0.497	0.652

Table 3.6: 95% credible interval bounds of correlations in individual-specific latent Fear, INR, and Quiz variation across draws from posterior

Variables	2.5% Quantile	97.5% Quantile
INR, Fear	0.028	0.127
Quiz, Fear	0.026	0.237
Quiz, INR	0.306	0.516

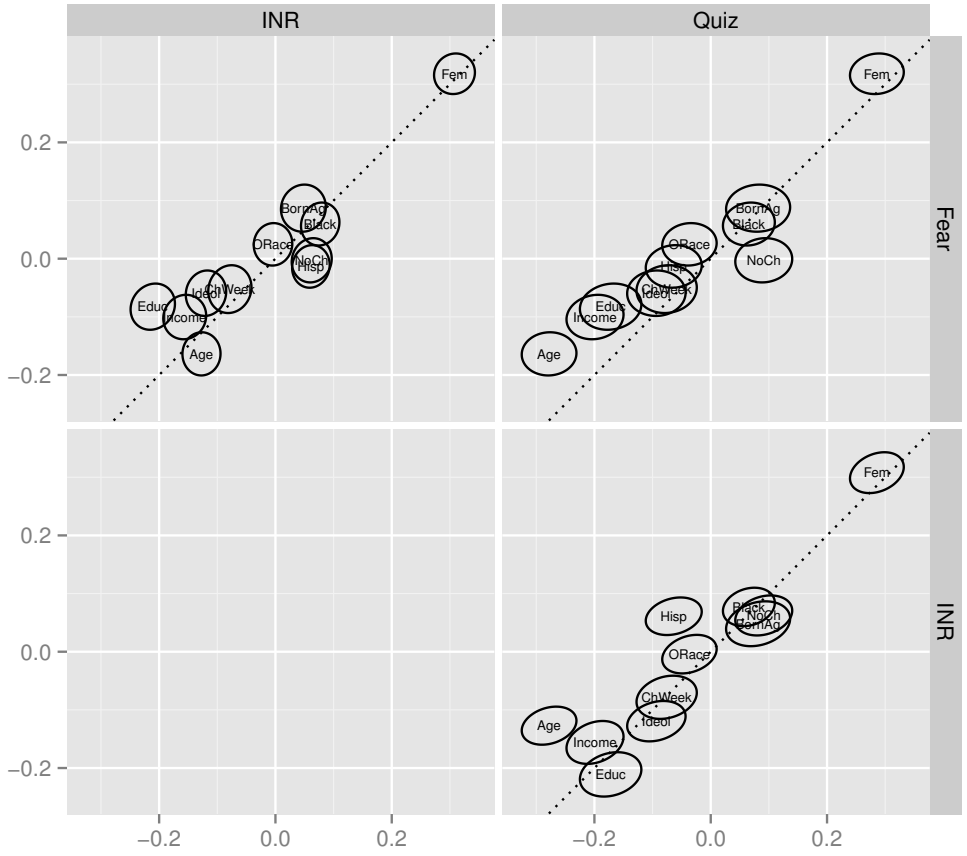


Figure 3.16: Fitted parameters of the Fear, INR, and Quiz model. The circles are the joint 40% confidence ellipses for each parameter in each pair of latent dimensions. 40% ellipses correspond to $t = 1$, and they are used to avoid excessive overlap in the visualization. 90% confidence ellipses are about 2.1 times larger, and 95% ellipses are about 2.4 times larger. The ellipses are based on the mean and covariance of each parameter's draws, which is an approximation of the actual joint distribution of the draws.

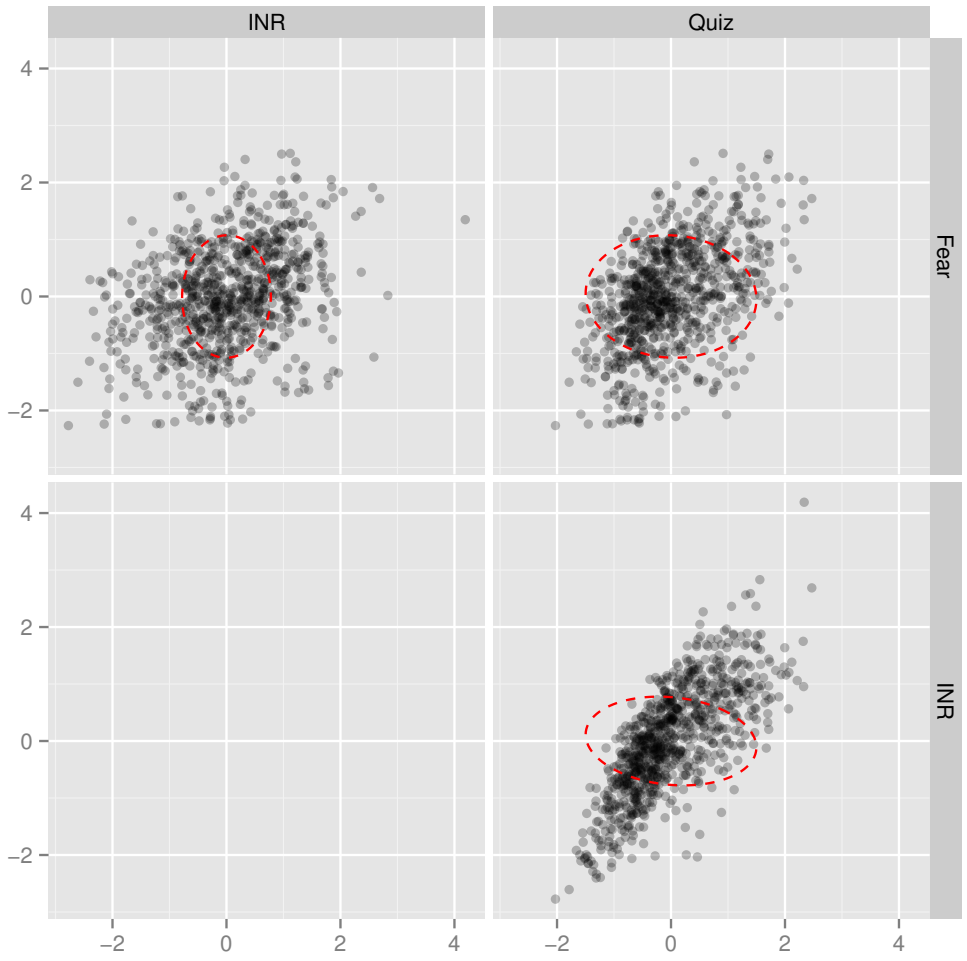


Figure 3.17: Mean latent Fear, INR, and Quiz estimates for each individual. The dotted red ellipse is a representative 95% joint credible ellipse for an individual at $(0, 0)$.

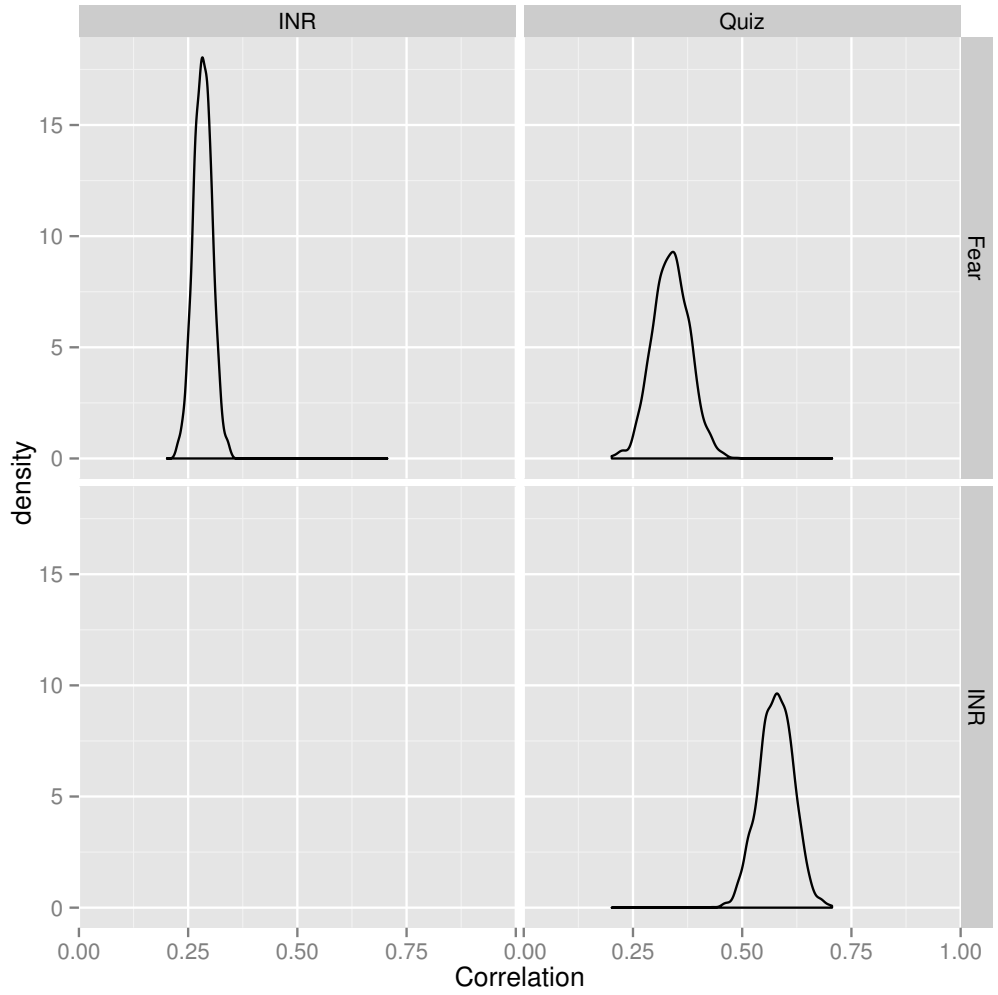


Figure 3.18: Distribution of correlations in latent Fear, INR, and Quiz parameters across draws from posterior

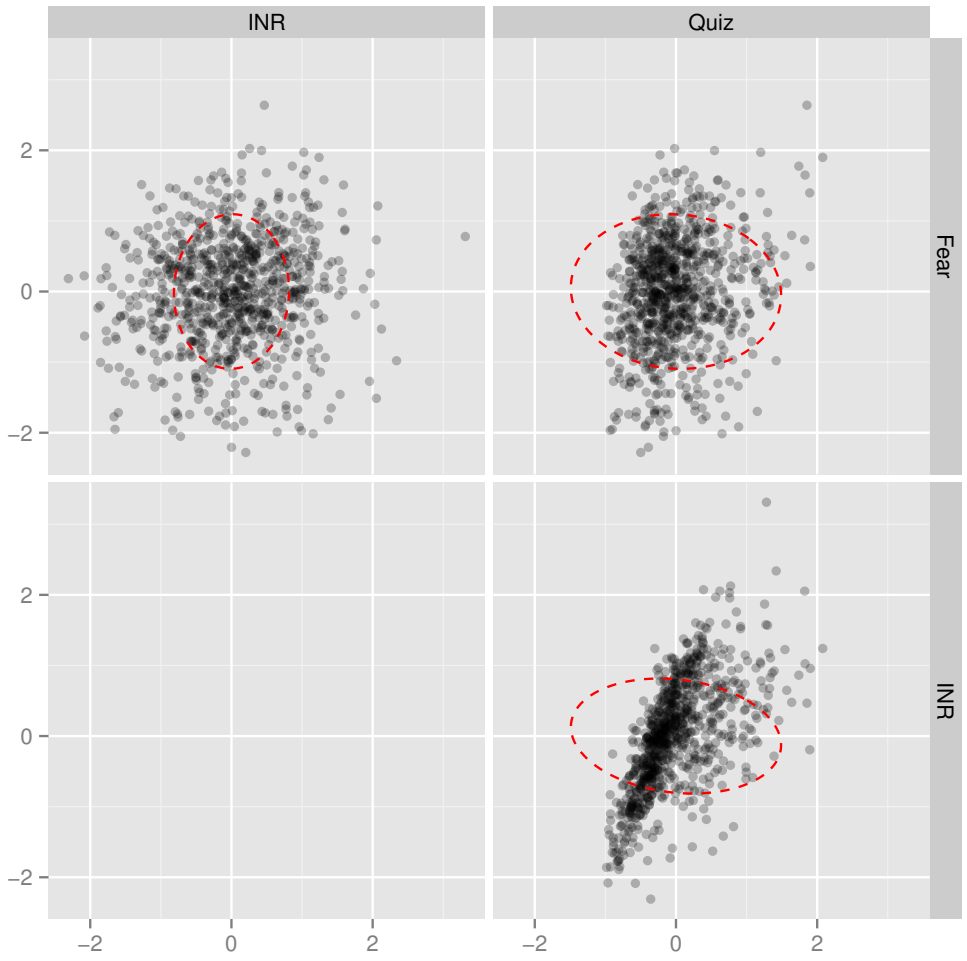


Figure 3.19: Mean individual-specific variation in latent Fear, INR, and Quiz terms. The dotted red ellipse is a representative 95% joint credible ellipse for an individual at $(0, 0)$.

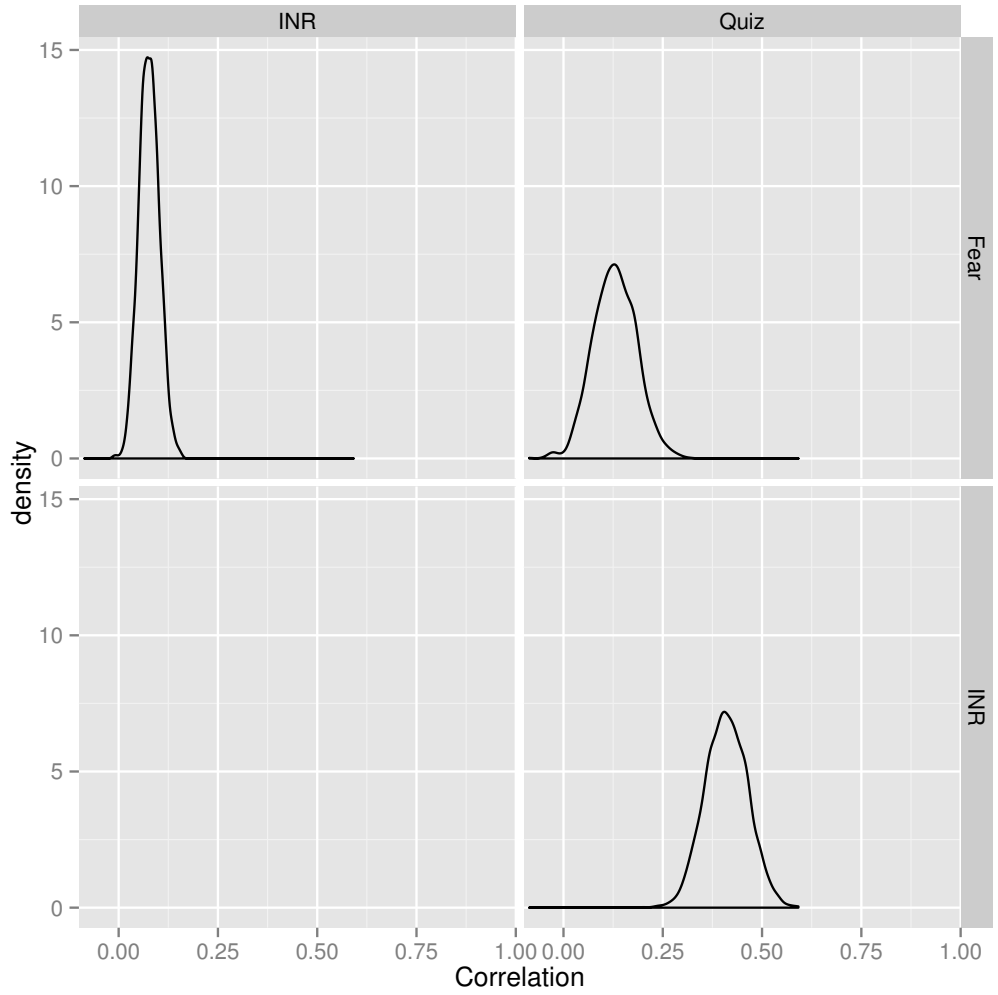


Figure 3.20: Distribution of correlations in individual-specific latent Fear, INR, and Quiz variation across draws from posterior

3.3.4 Willingness to admit

The willingness-to-admit model adjust the Fear/INR/Quiz model just discussed to account for the objective nature of Quiz, and the fact that factual knowledge has a very specific impact on item nonresponse. Political knowledge determines whether someone can create an answer they are comfortable with, but biases in willingness to admit that they do not have a good answer will change their 'goodness' threshold where they choose to skip a question.

This model is more similar to the traditional model discussed earlier in the paper, except that quiz questions are used rather than a political activity scale. Also, this model handles missing responses to questions cleanly, so it is possible to run against all the available data rather than just people who participated in both rounds of the survey. The traditional model showed varying effects between genders, so we try to capture those in this model as well.

The willingness-to-admit model uses latent factor variables for Quiz and Fear (similar to the unified model above, but replacing INR with Quiz), and then it uses those Quiz and Fear variables to predict the position on INR according to this specification:

$$(\text{Fear}, \text{Quiz})_s \sim \text{MVN}(Z_s\beta, \Sigma_s) \tag{3.8}$$

$$\begin{aligned} \text{INR}_s \sim\ & \beta_1 \text{Fear}_s + \beta_2 \text{Quiz}_s + \beta_3 \text{Fear}_s \text{Quiz}_s \\ & + \beta_4 \text{Fear}_s \text{Gender}_2 + \beta_5 \text{Quiz}_s \text{Gender}_2 \\ & + \beta_6 \text{Fear}_s \text{Quiz}_s \text{Gender}_2 \\ & + \epsilon, \qquad \epsilon \sim N(0, 1) \end{aligned} \tag{3.9}$$

along with ordinal logit constructions linking Fear, INR, and Quiz to their respective observable outcomes.

INR does not have demographic predictors in this model, so only the Fear and Quiz parameters are shown in Figure 3.21. There is less consistency in the

parameter estimate of Fear and Quiz than we saw in the previous section, and this is likely because Quiz is now forced to absorb some of INR's demographic variation. If we look at the Quiz/INR facet of Figure 3.16, Age and Hisp are more positive for INR than for Quiz, and when we compare the Quiz/Fear facet in Figure 3.16 with the Figure 3.21, we see that Age and Hisp were the main changes, and both moved in a positive direction. Whether this movement is a modeling artifact or a more accurate representation of political knowledge cannot be determined from the available data, but the changes are relatively small and so that distinction can be ignored with acceptable risk of misinterpretation.

The other plots of latent variables and their correlations are similar to the previous model, but with slightly more correlation between INR and the other terms that follow from the explicit structural connection. The main change in the plots of individual-specific variations and their correlations is that INR's noise is no longer correlated with Fear's or Quiz's. This is by construction since the individual-specific variation on INR is fixed to be independent of the other terms in Equation 3.9.

Table 3.7: Posterior quantiles of correlations in latent Fear, INR, and Quiz parameters across draws from posterior in the willingness-to-admit model

	Quantile				
Variables	2.5%	5%	50%	95%	97.5%
INR, Fear	0.260	0.267	0.303	0.341	0.347
Quiz, Fear	0.217	0.230	0.311	0.390	0.406
Quiz, INR	0.660	0.676	0.722	0.763	0.769

The key results of interest from the willingness-to-admit model are the effects of Fear and Quiz on INR from Equation 3.9. Distribution of the raw variables are shown in Figure 3.26 with quantiles in Table 3.9. The connection to Quiz is apparent, and a change of 1 in Quiz moves INR by nearly .75. The directions of other parameters are fairly uncertain, but we will discuss what they suggest. Fear

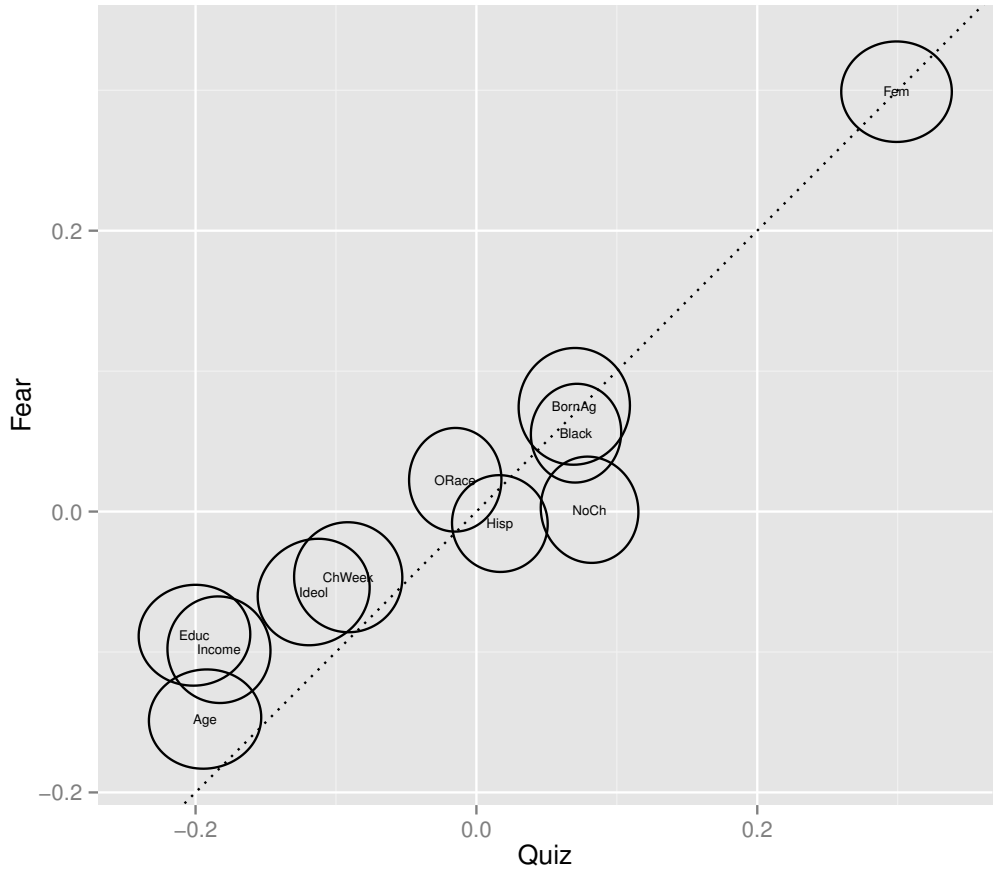


Figure 3.21: Fitted parameters of the willingness-to-admit model. The circles are the joint 40% confidence ellipses for each parameter in each pair of latent dimensions. 40% ellipses correspond to $t = 1$, and they are used to avoid excessive overlap in the visualization. 90% confidence ellipses are about 2.1 times larger, and 95% ellipses are about 2.4 times larger. The ellipses are based on the mean and covariance of each parameter's draws, which is an approximation of the actual joint distribution of the draws.

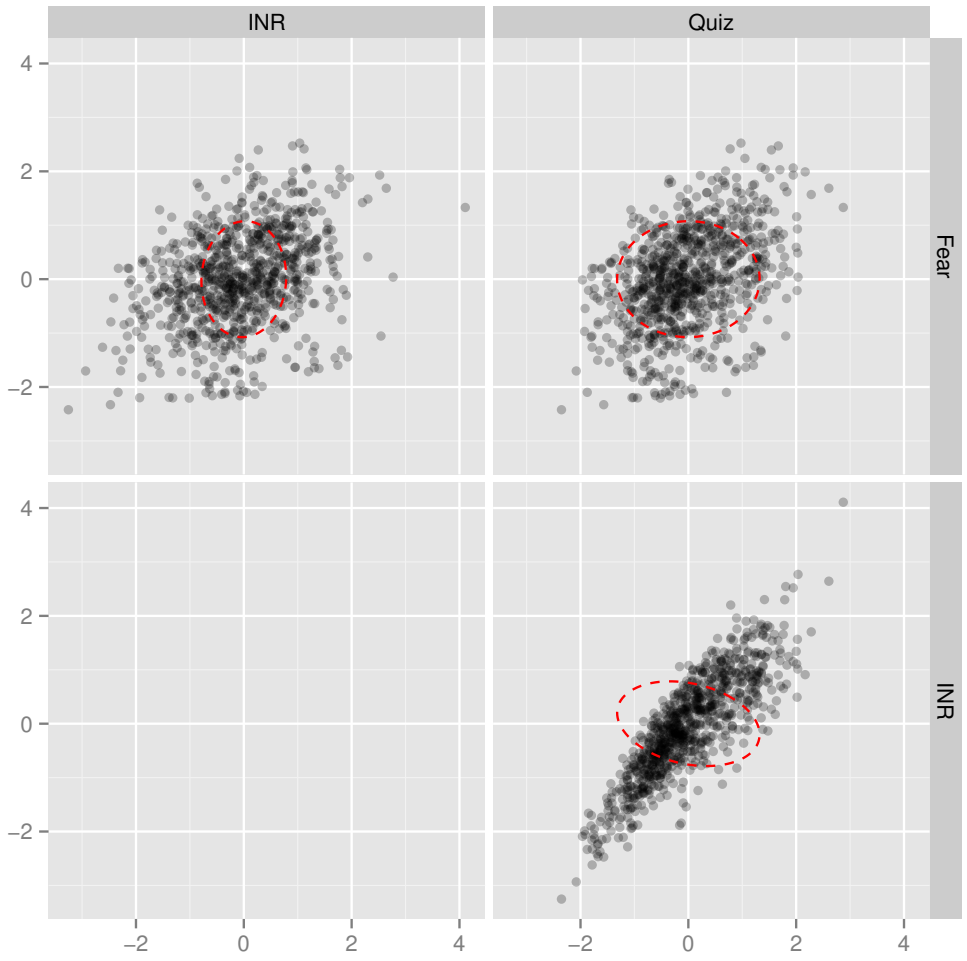


Figure 3.22: Mean latent Fear, INR, and Quiz estimates for each individual in the willingness-to-admit model. The dotted red ellipse is a representative 95% joint credible ellipse for an individual at $(0, 0)$.

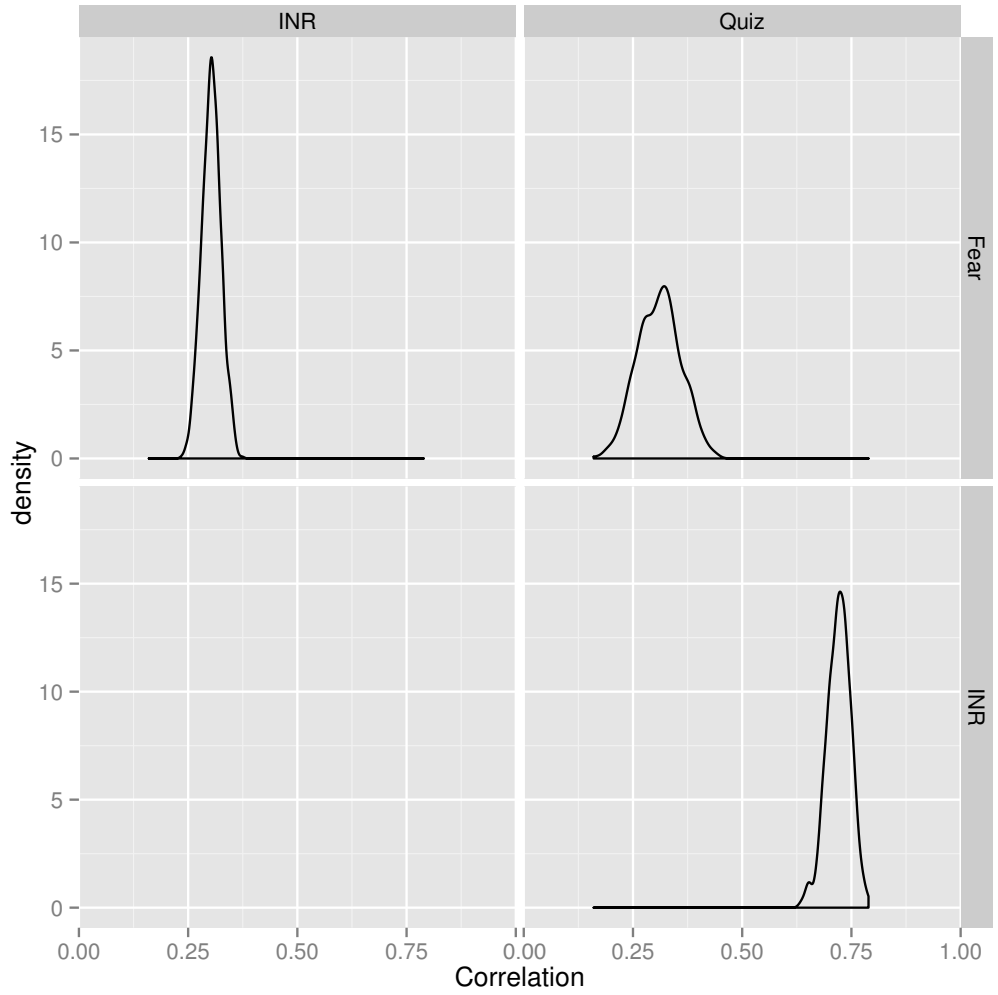


Figure 3.23: Distribution of correlations in latent Fear, INR, and Quiz parameters across draws from posterior in the willingness-to-admit model

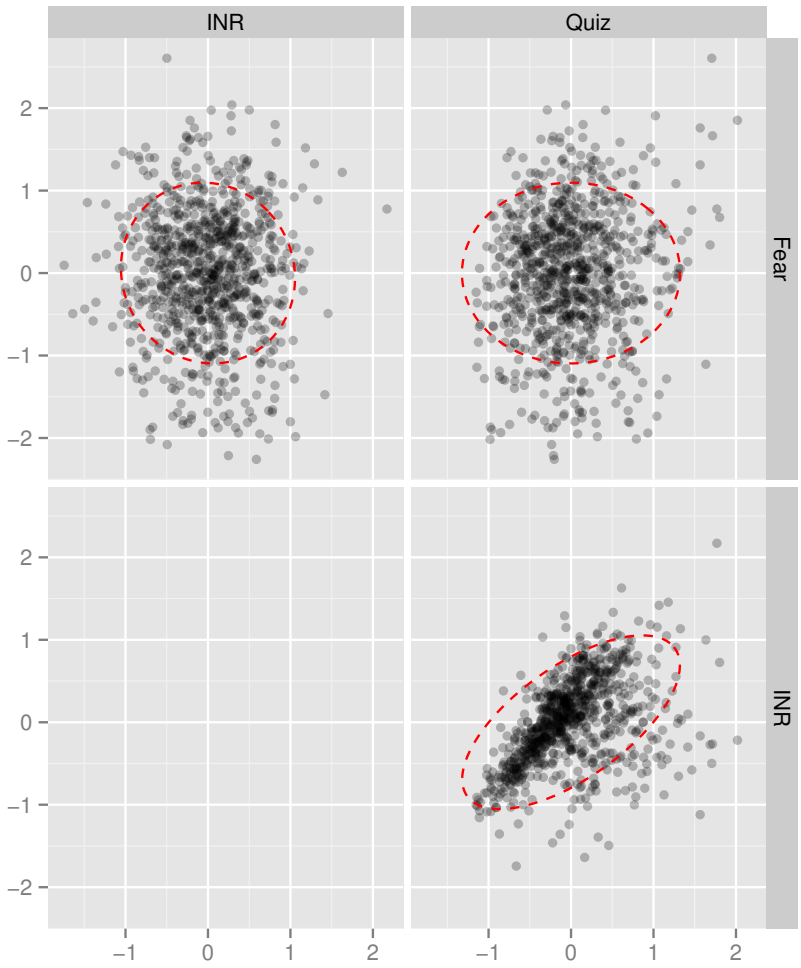


Figure 3.24: Mean individual-specific variation in latent Fear, INR, and Quiz terms in the willingness-to-admit model. The dotted red ellipse is a representative 95% joint credible ellipse for an individual at $(0, 0)$.

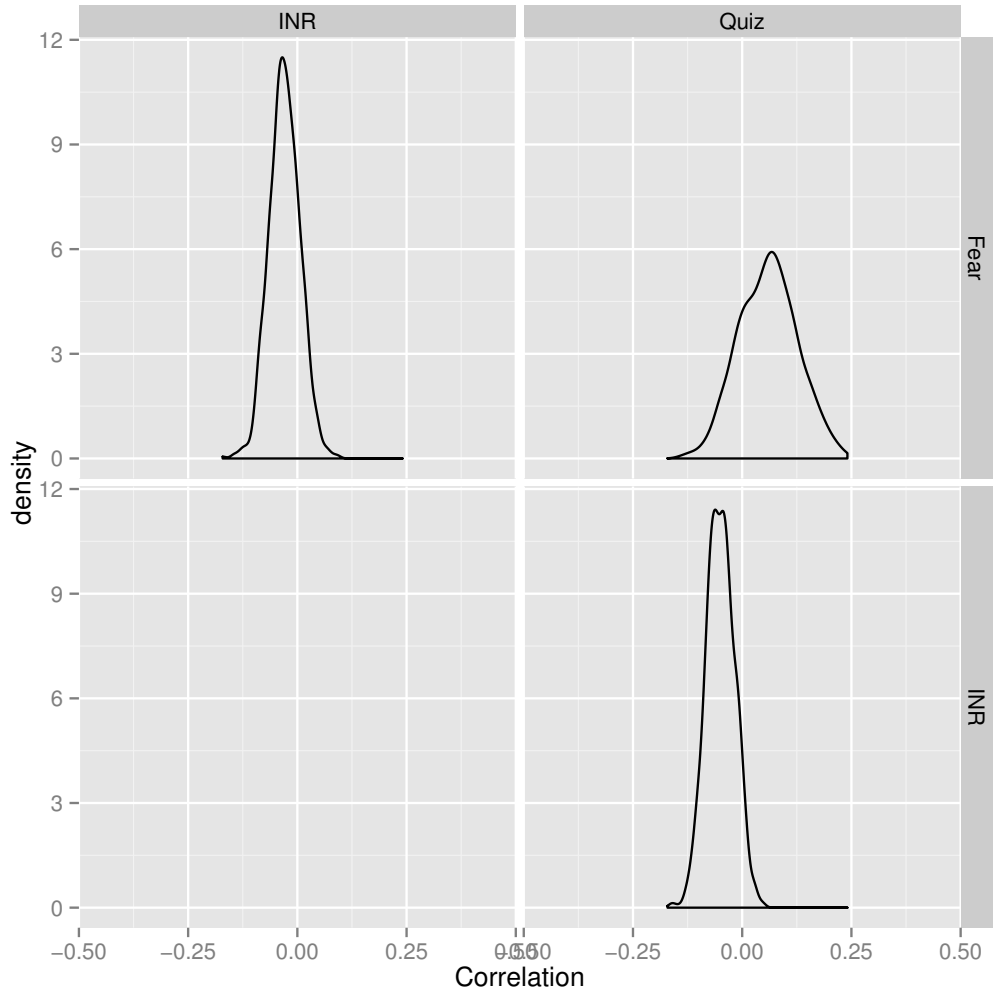


Figure 3.25: Distribution of correlations in individual-specific latent Fear, INR, and Quiz variation across draws from posterior in the willingness-to-admit model

Table 3.8: Posterior quantiles of correlations in individual-specific latent Fear, INR, and Quiz variation across draws from posterior in the willingness-to-admit model

	Quantile				
Variables	2.5%	5%	50%	95%	97.5%
INR, Fear	-0.094	-0.086	-0.030	0.026	0.041
Quiz, Fear	-0.067	-0.050	0.061	0.170	0.189
Quiz, INR	-0.112	-0.102	-0.050	0.003	0.011

appears to have a small positive effect on INR (more fearful people skipping more questions), and this effect appears to be strongest for people with less information (Fear $\times$ Quiz likely has a negative effect). There appears to be gender variation in this interaction, which will be clearer below.

Table 3.9: Posterior quantiles of Fear and Quiz effects on INR in the willingness-to-admit model

	Quantile				
Parameter	2.5%	5%	50%	95%	97.5%
Fear	-0.075	-0.038	0.096	0.228	0.249
Fear x Fem	-0.140	-0.116	-0.008	0.094	0.119
Quiz	0.622	0.647	0.766	0.911	0.938
Quiz x Fem	-0.128	-0.112	-0.027	0.054	0.069
Fear x Quiz	-0.184	-0.163	-0.068	0.011	0.025
Fear x Quiz x Fem	-0.044	-0.029	0.053	0.149	0.165

Comparisons between males and females are difficult in the graph with raw parameters, so Figure 3.27 and Table 3.10 transform the variables into separate effects for males and females. The impact of Quiz is apparent for both, and Fear's effect has a similar distribution for males and females, but the interaction between

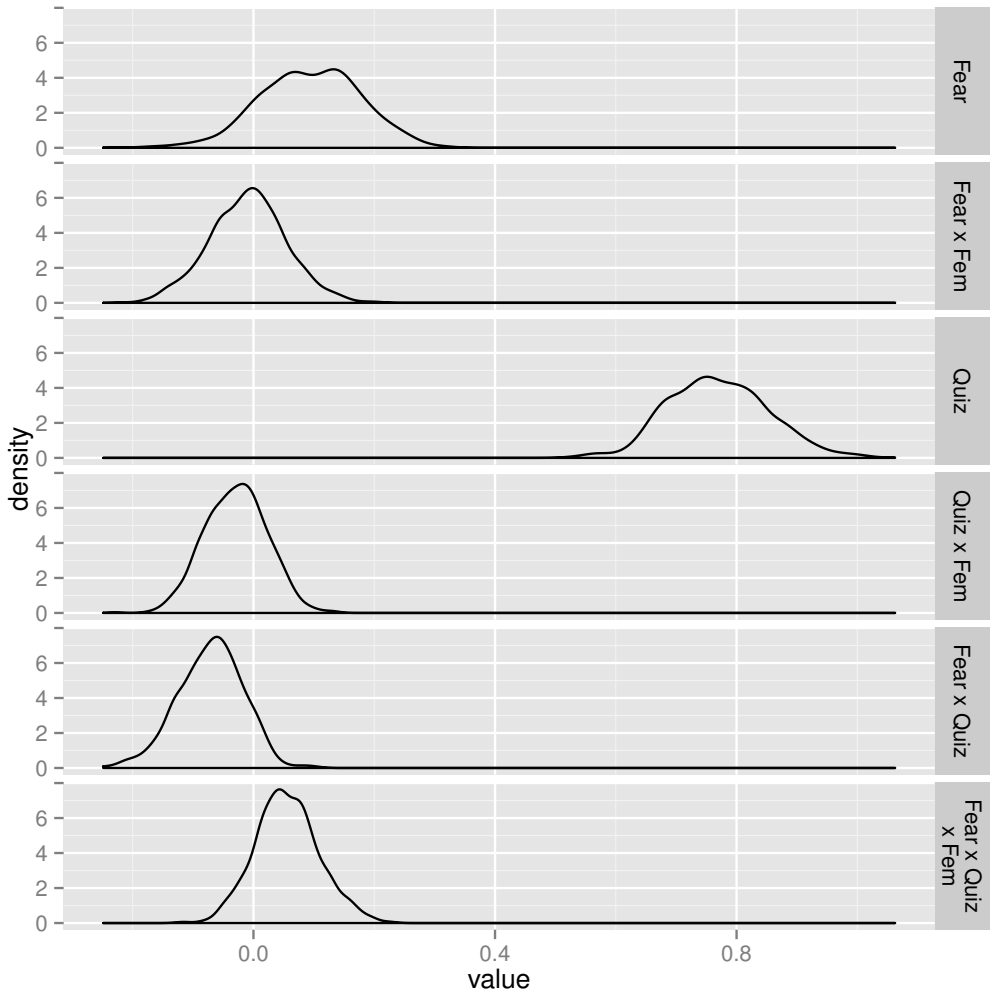


Figure 3.26: Distributions of Fear and Quiz effects on INR in the willingness-to-admit model

Fear and Quiz seems to only impact males—the estimate for females in noisy but has mean close to 0.

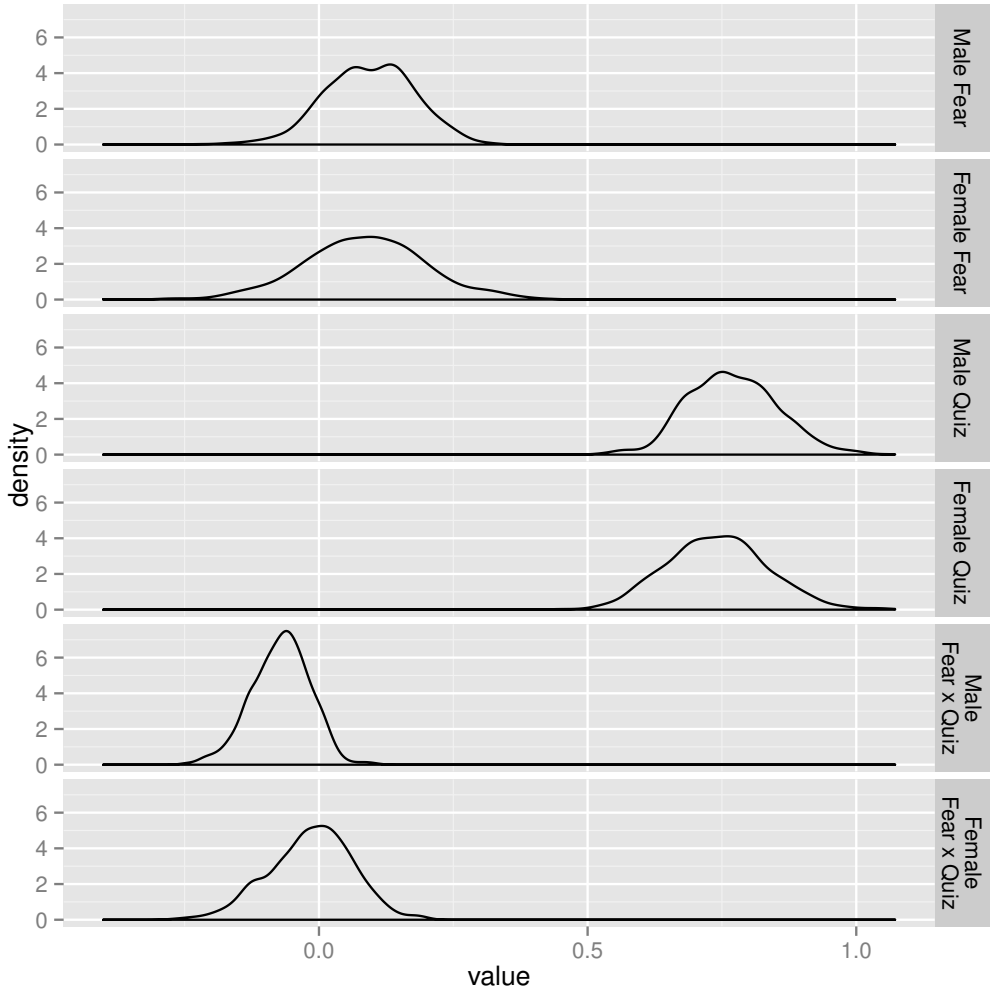


Figure 3.27: Distributions of Fear and Quiz effects on INR in the willingness-to-admit model transformed for easier comparisons between males and females

As mentioned before, only Quiz has a clear enough effect to be conclusive, but it is comforting to see that the interaction that we observed in the traditional model is reasonably well replicated in a model that is extremely different methodologically.

Table 3.10: Posterior quantiles of Fear and Quiz effects on INR in the willingness-to-admit model transformed for easier comparisons between males and females

	Quantile				
Parameter	2.5%	5%	50%	95%	97.5%
Male Fear	-0.075	-0.038	0.096	0.228	0.249
Female Fear	-0.143	-0.106	0.087	0.278	0.315
Male Quiz	0.622	0.647	0.766	0.911	0.938
Female Quiz	0.572	0.591	0.740	0.897	0.926
Male Fear x Quiz	-0.184	-0.163	-0.068	0.011	0.025
Female Fear x Quiz	-0.175	-0.146	-0.011	0.103	0.124

3.4 Conclusions

This paper has covered a very wide range of methods for modeling intuitive constructs like "political knowledge", "self-reported fear" and "likelihood of skipping questions" that cannot be measured directly by a single question. The traditional analysis used aggregate scores—simple sums of response values—and then ran a bundle of nested regressions to build up a story about how fear and item nonresponse are related. It had the benefit of methodological simplicity and familiarity for readers, but the results were tainted by well-known problems with aggregate scales.

The latent variable models started from a baseline of IRT modeling, which most methodologically inclined political scientists have heard of but are not necessarily familiar with. Despite that, the IRT setup allows for more intuitive models that operate directly on the latent terms that we are interested in—"political knowledge", "self-reported fear" and "likelihood of skipping questions"—and so the discussion can operate on a higher level separate from the complexity and noise of the ordinal IRT models below.

The latent construction had one unexpected benefit—it found similarities be-

tween demographic predictors of Fear and Quiz that would almost certainly be overlooked by a traditional analysis. It also clarified the connection between INR and Quiz on individual variation separate from demographics, and pointed to the possibility of omitted variables.

The biggest problem that remained for the latent variable models was the noise in estimating an individual's position on the Quiz dimension—uncertainty in an individual's estimation blurs them over nearly half the range of mean estimates. Unfortunately, factual questions are hard to come by in many survey data sets, including this one, and while the Quiz variable could probably be made more accurate by including participation terms as either predictors or as outcomes on the Quiz dimension, adding those terms would break the clean "political knowledge" interpretation and replace it with a fuzzier combination of political knowledge, political interest, and possibly network effects. The traditional model used a traditional workaround of just picking political activity as a proxy for knowledge, while the latent variable analysis used an 'enlightened' (or ignorant, depending on your perspective) approach of accepting the measurement uncertainty and letting the data say what it can on its own.

In the end, though, the general findings were reasonably similar between the traditional and latent variable models. Incorrect answers to factual questions and low activity both increased the likelihood of item nonresponse, fear appeared to increase INR rates, and there was an interaction between fear, quiz questions (or activity), and gender. Both modeling strategies were hindered by limitations of the data, though each was hindered in different ways, and while neither of the modeling approaches produced particularly strong results, the robustness of the qualitative interpretations to modeling choices is comforting for the development of political science. Just as ensembles of predictive models usually perform better than any single model, explanatory analyses can benefit when we use the full variety of analysis tools at our disposal.

Chapter 4

The connection between political and moral consistency

Our understanding of individual-level political behavior is rooted in notions of consistency. We ask how consistently group identity, values, wealth, or social connections are able to predict candidate choices, survey responses, campaign contributions, political interest, or turnout. Whenever we fit models of behavior or say whether one variable affects another—basically, whenever we engage in quantitative political analysis at the level of individuals—we are talking about consistency.

This is particularly true when we try to understand political values. Values are not directly observable and we are unable to reliably measure them in surveys, so we instead have to look for patterns across multiple behaviors that could be explained by some underlying trait, value or belief. The dominant tools tend to project many observed outcomes onto some low-dimensional space of latent variables, either through item response theory models, factor analysis, self-reports, or simple aggregations (Feldman, 1988; Gerber *et al.*, 2010; Kanai *et al.*, 2011; Sniderman and Bullock, 2004; Zaller and Feldman, 1992). We sometimes approach the problem with weak hypotheses about what those latent variables or groups might be (in the case of item response theory or unrestricted latent class analysis), and at other times, we have strong a priori hypotheses about the specific variables in question and seek to test whether the data appears consistent with those hypotheses (in the case of aggregations and self-reports). The field has focused on

these "exploratory vs. confirmatory" questions, discussions about the best linkages between the latent variables, and analyses of out of sample performance.

Political scientists used to be much more interested in questions of individual consistency with models—such questions were the focus of *The American Voter* (Campbell *et al.*, 1960) and Converse (1964)—but the quantitative literature on estimation of political values never really caught up with those comparatively qualitative analyses. Through the long development of ideal point and spatial preference models, we have been placing everyone on an identical latent space.[1] Intuition points in the opposite direction; people vary immensely in how they parse political questions. Maybe they watch different news outlets and latch on to different key words in a question or the questions remind them of different experiences in their past. We treat this inconsistency as a problem to be addressed through better survey questions or more robust modeling techniques. The spatial ideology literature does not consider that people vary in their individual adherence to their underlying values when making choices—or they just have a different set of underlying value dimensions.

Inconsistency has not always been pushed aside as an annoyance, though. Ever since Converse (1964) and *The American Voter* (Campbell *et al.*, 1960) found that people do not explicitly structure their choices around abstract ideologies, researchers have been trying to explain why, then, people do not just behave randomly. Political scientists set out to understand what factors, often education or political interest, make people behave consistently over time or in predictable ways based on demographics or stated beliefs. This line of research tended to look at one specific policy at a time, and it never examined patterns of individual consistency that span multiple polices or that link between consistency of policy choices, values, and behaviors. It also never examined possible links between consistency in the political domain and consistency in other areas that could shed light on potential mechanisms by which the consistency is created.

[1] Latent class analysis is somewhat less restrictive, but still assumes homogeneity in the goodness-of-fit within groups.

In terms of methodology, political scientists are able to estimate latent policy dimensions, and we can characterize consistency on specific policies, but we are still far from the ideal of simultaneously characterizing both the latent policy dimensions and the variation in individual adherence to those dimensions. This chapter attempts to bring us a bit closer to that ideal. I adapt models from the legislative spatial voting literature to simultaneously estimate individual consistency with a multidimensional spatial voting model and the dimensions underlying that model. 'Consistency' in this model takes on multiple dimensions as well, which allows people to behave in accordance with some, all, or none of the underlying values. Unifying our models of consistency and our models of political values can give us greater insights into the structure of political thoughts and decisions than either approach alone. I also compare consistency on political dimensions with consistency on a set of moral dimensions to clarify the extent to which political consistency is unique or just a facet of a more general psychological phenomenon.

4.1 Consistency

People like to be consistent in how they think and answer questions, or they at least like to view themselves as consistent people. In the realm of politics, this could lead them to pick news outlets that agree with their existing views (Iyengar *et al.*, 2008) or to change their true or stated beliefs to fit group identification or vice versa (Bafumi, 2009; Gerber *et al.*, 2010), and generally adjust their answers and behaviors to follow some sort of discernible structure.

But what is "consistent" varies from person to person. Voters vary in their desire and ability to be consistent, or to be perceived as consistent by any particular observer, and political elites and highly informed or active voters very likely think about the structure of politics differently and therefore diverge in what it means to be consistent. From this angle, it is not particularly surprising that Converse (1964) and *The American Voter* (Campbell *et al.*, 1960) found that most people do not explicitly structure their political ideas and choices around abstract philo-

sophical ideals, with the general conclusion being that they rely on group identity instead. But, as the decades since have shown, abstract ideology and group identity are not the only reasons that voters might end up with a 'consistent' set of policy and voting choices.

Zaller and Feldman (1992) and others have demonstrated that a voter's consistency tends to increase with their political informedness. Zaller models the voter's decision as a the result of a weighted sum of randomly sampled 'considerations' that may push a voter for or against a policy. Additional information tends to increase the number and concordance of considerations that the voter samples in making their choice, and so they tend to make more consistent choices.

Alvarez and Brehm (1998, 2002) argued for an alternative model that incorporated underlying values or beliefs into the voter's choice. Under perfect information, specific policy questions relate to one or more of these underlying values, but if the relevant values support conflicting answers (ambivalence), additional information is unlikely to reduce a voter's variance. If multiple values support the same choice (equivocation), then additional information will simply make the voter's choice clearer, and if the voter simply has little information about the policy (uncertainty), then additional information will reduce a respondent's variability, but not in ways determined by underlying values. (Alvarez and Brehm, 1998) predicted whether an individual agreed or disagreed with statements about the IRS (the employees are honest/knowledgeable, the IRS is too intrusive, they never try to take more money than they should) given opinion metrics based on answers to other questions about the tax system, the respondent's interaction with the IRS, and various demographics. Alvarez and Brehm allowed for structured heteroskedasticity conditional on informedness, education, and interactions between the opinion metrics, and they did find clear evidence of variation in individual consistency, part of which was due to conflicting preferences, and part of which was related to informedness or education.

This previous research is more muddled than it might appear. Political scientists have a number of different meanings for voter 'consistency': the stability

of an individual's answers over time; the predictability of one policy preference conditional on other preferences; or the predictability of specific policy choices conditional on underlying values (Sniderman and Bullock, 2004). Since the third definition, predictability conditional on latent values, roughly encompasses the previous two, that is the type of consistency that I will address in this chapter.

With these multiple definitions in mind, Zaller's model is most powerful at explaining why some individuals have more stable survey responses over time than others: large repeated samplings from a homogenous pool will have low variance; but it has limited utility in explaining how consideration sets relate to underlying values. Alvarez and Brehm's work also addresses individual stability over time to a certain extent, since underlying values are assumed to be more stable than information. They attempt to address consistency around underlying values directly, but by modeling each policy separately, they are forced to use proxies for underlying values, and so they risk muddling the relationship between uncertainty, ambivalence, and equivocation with heterogeneity of values within proxy strata. Perhaps most importantly, none of these models get us closer to understanding what factors other than informedness might influence political consistency.

All empirical questions about voters' consistency with underlying values hinges on being able to estimate or characterize those values in some way. Since *The American Voter* clearly ruled out explicit conscious adherence to underlying values, the only way to get at these underlying values is to measure patterns of observable behavior. Thus, the main approaches taken in both the legislative voting and public opinion literature are various forms of factor analysis that either implicitly or explicitly rely on an spatial model to connect underlying values to specific policy choices. Poole and Rosenthal (1985) was the foundation for much of the legislative voting literature, and in the years since, they and others like Heckman and Snyder (1997) extended their models to include dynamic changes in values, issue-specific characteristics, and to a certain extent, heteroskedasticity. Jackman (2001) and Clinton *et al.* (2004) moved the spatial model to a Bayesian framework based on multidimensional item-response theory models from psychometrics, which made it

much simpler to deal with the large number of latent parameters inherent in factor analysis problems.

Very recently, Lauderdale (2010) extended the Clinton *et al.* (2004) model to allow for legislator heteroskedasticity. He terms high-variance legislators 'mavericks', but they could as easily be called 'inconsistent' under the earlier definition since such legislators do not reliably follow the same set of underlying latent values that the others do.

In the remainder of the chapter, I demonstrate that Lauderdale's legislative estimator can be extended further to model the consistency of survey respondents with their underlying latent values. I then compare respondents' consistency on a set of political questions with their consistency on a set of moral questions and find evidence of causes other than political informedness that influence how consistent a voter will be. I do not attempt to address what exactly causes variation in political and moral consistency, only that there is a relationship between the two.

4.2 Formal and statistical models

Political and moral decisions are both treated as outcomes from separate K-dimensional spatial where $K = 2$ for political decision and $K = 5$ for moral decisions. Each individual s has an ideal point $\phi_s = (\phi_{s1}, \ldots, \phi_{sK})$ in policy space, and each question q has a policy direction vector $\lambda_q = (\lambda_{q1}, \ldots, \lambda_{qK})$, η_q possible answers, and a set of cutpoints $\zeta_q = (\zeta_{q1}, \ldots, \zeta_{q\eta_q})$ that define the breaks between answers.

I use a simple ordinal logistic model for answers to questions. Individuals map their ideal points ϕ_s onto the question vector λ_q, then add idiosyncratic noise ϵ_{sq} to get x^*_{sq}, their ideal answer to the question, and finally they choose y_{sq}, the closest available answer to their ideal answer. Since the questions are all ordinal, we can define a set of cutpoints that give the midpoints between successive answers, so for a question with η_q possible answers, we get $\eta_q - 1$ cutpoints that define η_q acceptance intervals. We can thus equivalently say that an individual finds the

acceptance interval that contains their ideal answer, and they then choose the corresponding available answer.

Formally, this can be written

$$x^*_{sq} = \lambda_q \cdot \phi_s + \epsilon_{sq} \tag{4.1}$$

$$y_{sq} = \begin{cases} 1 & , \quad x^*_{sq} < \zeta_q 1 \\ \vdots & \\ i & , \quad \zeta_{q,i-1} \leq x^*_{sq} < \zeta_{q,i} \\ \vdots & \\ \eta_q & , \quad \zeta_{q\eta_q} \leq x^*_{sq} \end{cases} \tag{4.2}$$

where ϕ_s does not include a constant term.

The arrangement so far is very similar to those used in legislative voting analyses like Clinton *et al.* (2004) except that this version allows for ordinal rather than only binary decisions.

The key difference comes with how ϵ_{sq} is parameterized. I allow for heteroskedasticity across individuals in the idiosyncratic noise terms:

$$\epsilon_{sq} \sim \text{Logistic}(0, \sigma_s^{g(q)}) \qquad \text{iid over } q,\ s \tag{4.3}$$

$$(\sigma_s^p, \sigma_s^m) \sim \text{LogNormal}(Z\gamma, \Sigma^\sigma) \tag{4.4}$$

where $g(q)$ denotes whether question q is a political or moral question, and σ^p and σ^m are the political and moral heteroskedasticity terms. $Z = (z_1, ...z_D)$ is a matrix of demographic predictors where each column has been centered to have mean 0 (and in practice, scaled to have standard deviation 1). Compared to models with question-specific heteroskedasticity, this arrangement greatly reduces the dimensionality of the heteroskedasticity term and aids general interpretation of overall consistency rather than question-specific consistency. Note that this arrangement is very similar to the heteroskedastic one-dimensional legislative voting model of Lauderdale (2010) but applied to ordinal data and extended to multiple dimensions

of preferences and heteroskedasticity.

Finally, ϕ_s and Σ_s are influenced in the model by hierarchical covariates:

$$\phi_s \sim \text{MVN}(Z\beta, \Sigma^\phi) \qquad \text{iid over } s \tag{4.5}$$

$$\Sigma^\phi \sim \text{Wishart}(I, 14) \tag{4.6}$$

where Σ_s allows for covariance between the political and moral dimensions, and the degrees of freedom parameter is twice the total dimensionality of the latent space and provides moderate shrinkage towards the identity matrix.[2] The value of any particular ϕ_s is not of much interest to us, as it's just the estimate for one specific individual, so they are effectively nuisance parameters in the model. As we will see later, we can differentiate some individuals' values of ϕ_s from others, but the real interest is in the group-level values of β and Σ^ϕ.

Identification of latent ideal point models depends on fixing some of the latent parameters to define the scaling and rotation of the latent space (Clinton *et al.*, 2004; Jackman, 2001; Rivers, 2003). Legislative ideal point models usually fix the latent space by fixing the ideal points of a few legislators, with the number depending on the policy space's dimensionality. That is not possible here since the data comes from a random sample of voters and there is no prior information about individual voter preferences.

Further complicating the matter is the fact that the moral questions were designed to have a factor structure that can be identified easily through constraints on question parameters but no orthogonality constraints in the latent space. The political questions, on the other hand, have no such predetermined structure and are easiest to interpret with orthogonal dimensions. The sampler can operate without strict orthogonality on the political space, but if the two political dimensions become too correlated, the sampler can hit regions of numerical instability. To avoid this, I allow correlation in the political dimensions within the sampler, but

[2]These parameters create a fairly sharp drop-off in the prior likelihood of correlations over 0.5 between dimensions.

I add a penalty term to such correlations (which hurts mixing slightly but does not bias the results), and then run the actual orthogonalizations after sampling is complete. This takes care of orthogonalization, but we still need to pin down the political dimensions' scale and rotation.

I fix the relative scaling of ζ and ϕ by standardizing the noise term of ϕ to mean 0 and identity variance. I also constrain the political questions to have weight 0 on all moral dimensions, and vice versa, and while the directions of moral questions are specified *a priori*, those of political questions are not. Instead, I run the sampler on a version of the model that does not identify the political space's rotation, and then I apply Procrustes rotations to the political ϕ drawn in each sample to align them to a common space across all samples. Finally, I manually rotate the political dimensions to simplify discussion.

The final step in identification is checking the cutpoints ζ. The scaling of ϕ is fixed to mean 0 and variance 1 on each dimension, so if we leave λ unconstrained, it is sufficient to check that the mean and scale of ϵ_{sq} are both fixed. The mean of ϵ_{sq} is constrained since it is normally distributed with mean 0, and the scale is constrained because the log standard deviation of ϵ_{sq} has mean 0 across individuals.

Note that the identification restrictions in latent parameters are distinct from assumptions in that it is possible to transform draws taken from the posterior distribution of this model into draws from the posterior of models with other identification restrictions. In other words, the identification restrictions should be thought of as definitions that make it possible to compare points in latent space, but not as informative restrictions on the model.

4.3 Data

All data used in this chapter comes from the 2010 Cooperative Congressional Election Survey (CCES), an internet survey run by Polimetrix over a representative sample of US voters. The CCES includes two rounds, one before the 2010 election and one after, and in each round, respondents answer a set of common questions

followed by a set of questions from a team module. There were 46 teams, each of which was assigned 1000 or more unique respondents. In this analysis, I use the results from the Caltech and New York University modules, which both included the battery of moral questions. All questions are ordinal or are transformed into ordinal variables, so they all fit immediately into the model presented above. The actual political questions are listed in Appendix A. Moral questions come from the 30-item Moral Foundations Questionnaire (Graham *et al.*, 2011), and each respondent was randomly assigned one of three subsets the questions in order to reduce the length of the survey.

4.4 Results

The results are broken into two sections. The first section fits a model without heteroskedasticity, and the second section fits a model with heteroskedasticity. While the second model is the main focus of the chapter, the first model provides a basis for comparing the fitted latent spaces. The purpose of this example is to demonstrate the performance of the estimation methodology on a real-world data set, and here it is a primarily descriptive exercise. With that in mind, I will not analyze the specific findings of the model very deeply. The discussion below will focus on the performance of the model for differentiating people and groups in ideological space and in their consistency, not on the specific impacts of any particular demographic.

4.4.1 Model without heteroskedasticity

The model without heteroskedasticity fixes the (σ_s^p, σ_s^m) terms to $(1, 1)$ for everyone but is otherwise identical to the heteroskedastic model developed above. The model was coded in Stan, and I started with a series of tuning steps to get the sampler into the 'meat' of the posterior, and then ran Stan on four chains with 50 adaptation steps and 200 samples each. The chain length is much lower than would be required for hierarchical IRT models that use Gibbs sampling or Metropolis-Hastings, but

Table 4.1: Demographic variable descriptions

Variable	Description
Unins	Respondent is uninsured
UnionHouse	Someone in household is or was a union member
UnionSelf	Respondent is or was a union member
Rent	Respondent rents their home
Income	Ordered income bins (treated as continuous)
Pray	Frequency of prayer
Church	Frequency of church attendance
Marr	Respondent is married
Educ	Education bins (treated as continuous)
RaceO	Race other than non-Hispanic White or Black
Hisp	Hispanic
Ret	Respondent is retired
NoFT	Respondent does not have a full-time job
Fem	Female
Black	Race is Black
BornAg	Born-again Christian

the Hamiltonian Monte Carlo algorithm that Stan uses walks across the posterior several orders of magnitude more efficiently, so 200 samples per chain produce good mixing.

There are three outputs of interest: the hierarchical effects of demographics on ideal points, distributions of each individual's ideal points, and the moral and political question loadings.

The goal of this chapter is not to dig into the predictors of political and moral opinions, so the main point to take from Figures 4.1 and 4.2 is that the model is does estimate the demographic predictors with enough precision to make meaningful group-level claims, and the demographic predictors fall onto a reasonable set

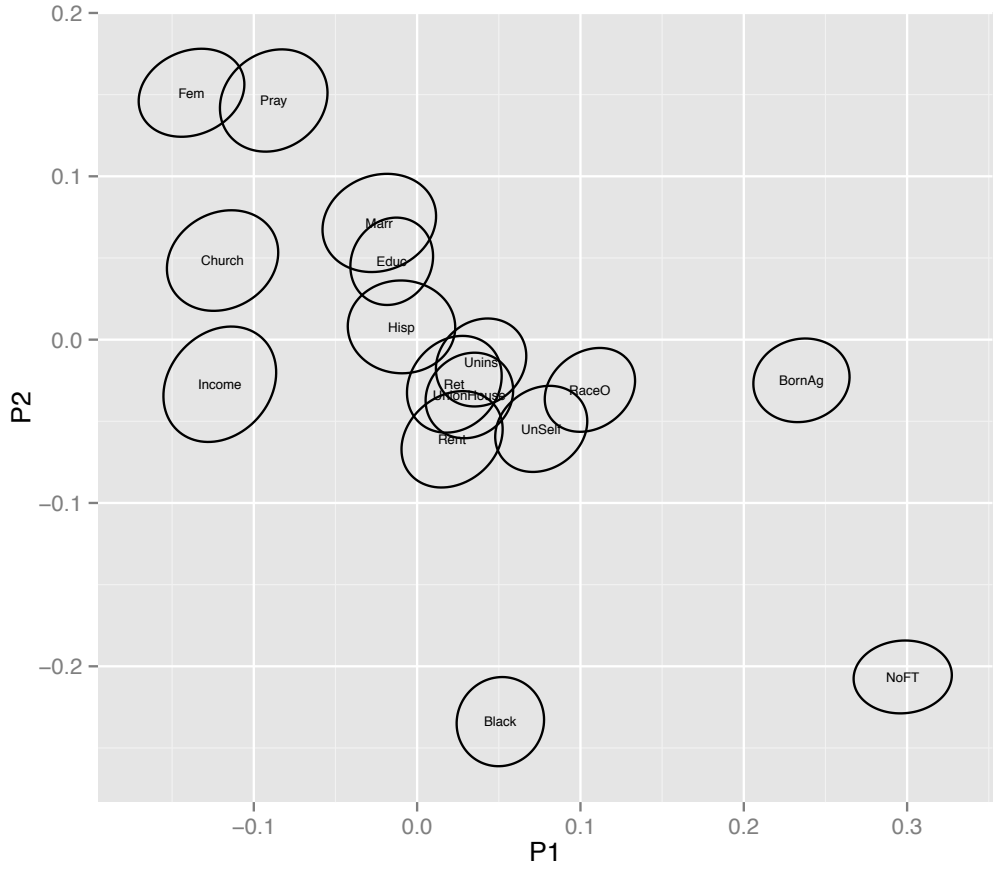


Figure 4.1: Effects of demographic parameters on political ideal points in model without heteroskedasticity. Ellipses are the 40% confidence ellipse for the effect estimates (roughly comparable to 1 standard deviation).

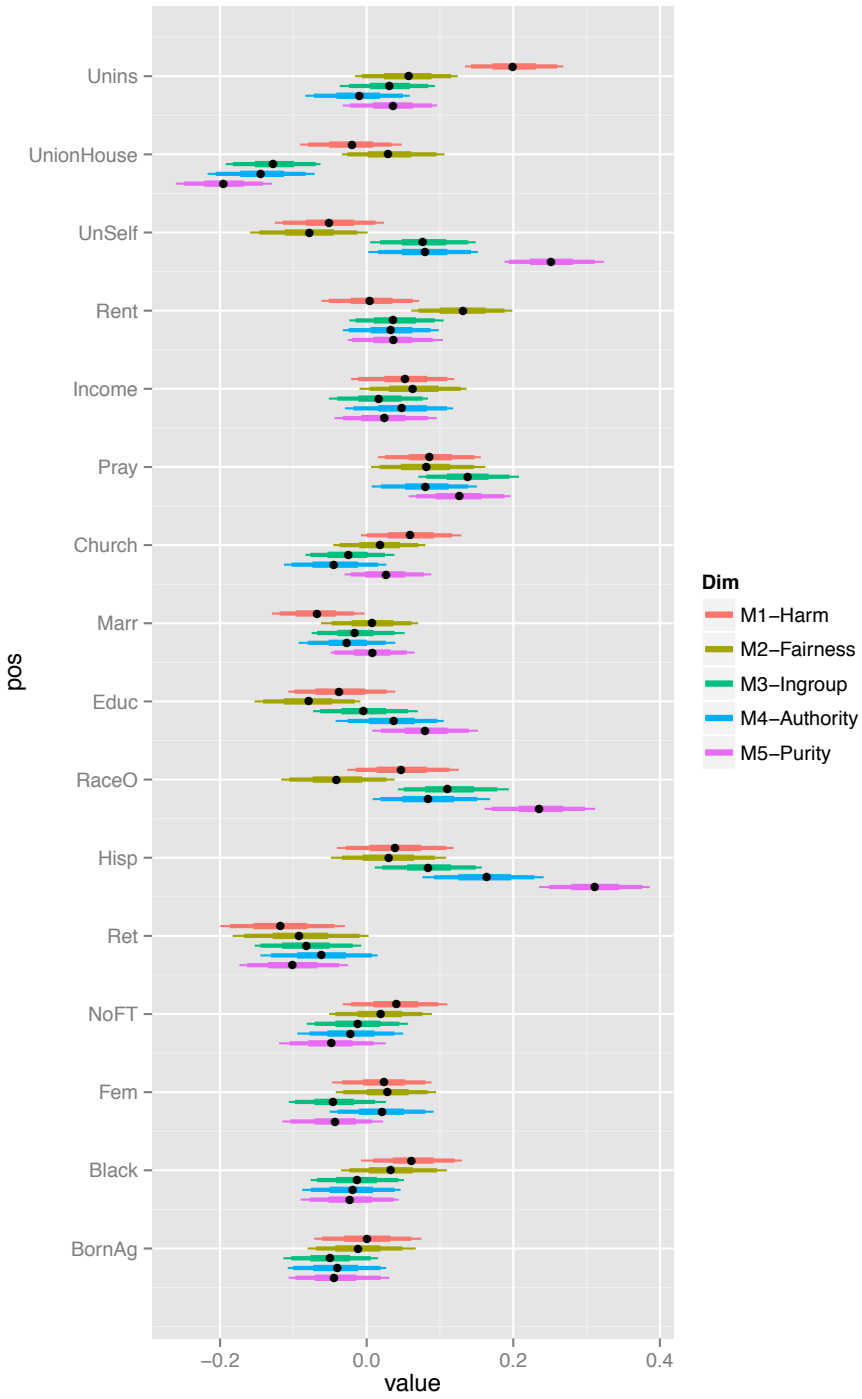


Figure 4.2: Effects of demographic parameters on moral ideal points in the model without heteroskedasticity. Error bars are stepped at 60%, 80%, and 95% levels.

of dimensions. Income and NoFT have effects on primarily the (P1−P2) diagonal, as do most of the demographics, while Church, Pray, and Black spread away from this diagonal. For descriptions of the variable definitions, see 4.1.

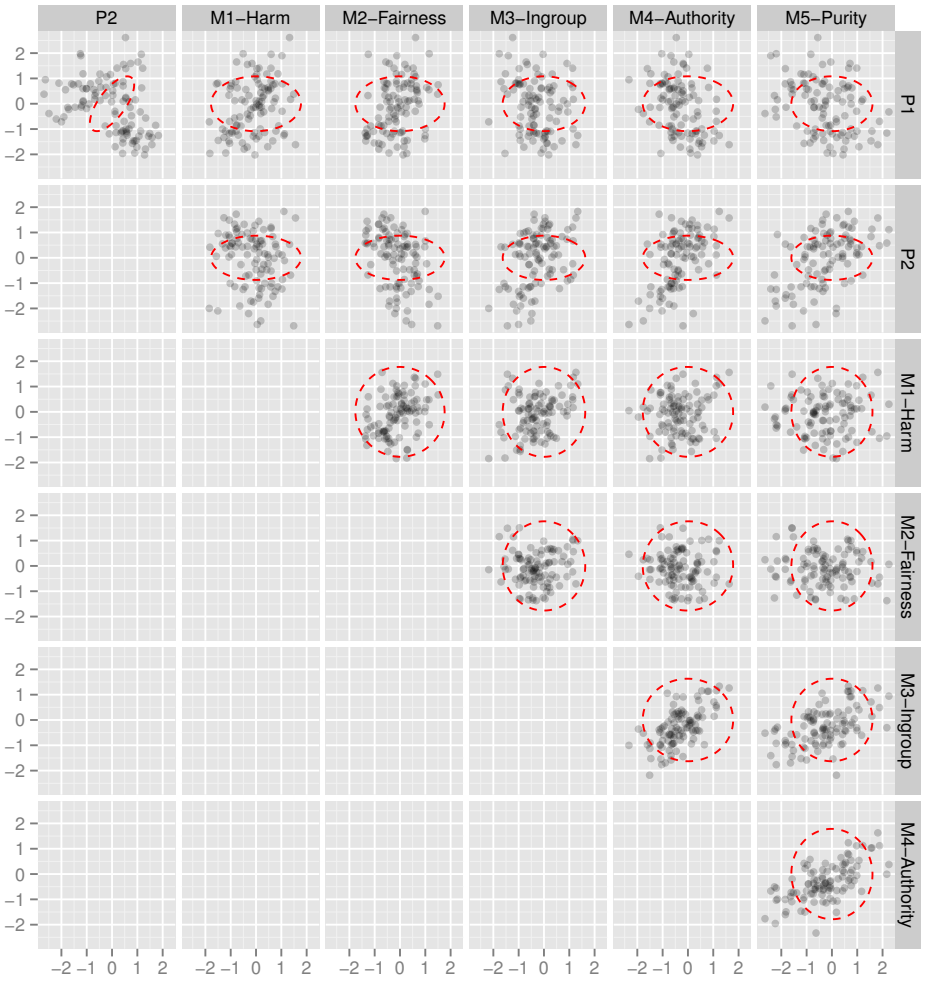


Figure 4.3: Scatterplots of a sample of the fitted ideal point in the model without heteroskedasticity. Red dotted ellipses denote the 95% confidence ellipse for the ideal point of a hypothetical individual at 0.

The plots of ideal points in Figure 4.3 show that the model is differentiating individuals reasonably well (note that the red circles are 95% ellipses, so people at opposite edges are quite well differentiated), but that there is relatively more

uncertainty on the moral than on the political dimensions when we try to estimate an individual's position.

Finally, the factor loadings in Figure 4.4 shows that we have a good distribution of political questions that are informative about different directions in the ideological space. The specific effects are not of interest here and with 196 variables, there are far too many to discuss. The mass of large, thin ellipses also shows that there are many questions that did not have reliable factor loadings (usually because they are only asked to small subsets of respondents), but there are enough well-estimated loadings to differentiate people well.

The moral loadings in Figure 4.5 split up the moral questions by their dimensions as defined by the MFQ30, though while the standard method of scoring the MFQ30 places equal weight on every question, we see a very wide variation in the loadings within each dimension. Again, this is not the focus of this chapter, so it will not be discussed in detail.

4.4.2 Heteroskedastic model

Like the homoskedastic model, the heteroskedastic model was coded in Stan, and I started with similar set of tuning steps, and then ran Stan on four chains with 50 adaptation steps and 200 samples each. The chains showed good mixing during the sampling stage.

There are a few substantial changes in the demographic predictors in Figure 4.6 compared to the homoskedastic version in Figure 4.1. NoFT now impacts P2 more negatively, and Church, Fem, and Pray have a more negative impact on P1. These movements will tend to be in the direction of low-heteroskedasticity individuals with high values of the parameter, so if the low-heteroskedasticity non-full-time workers are low on the P2 dimension, that will tend to pull the new estimate of NoFT downwards.

There are no substantial relative differences in the effects of demographics between the heteroskedastic model (Figure 4.7) and homoskedastic model (Figure 4.2). The heteroskedastic model deflates the scale of all the effects by roughly

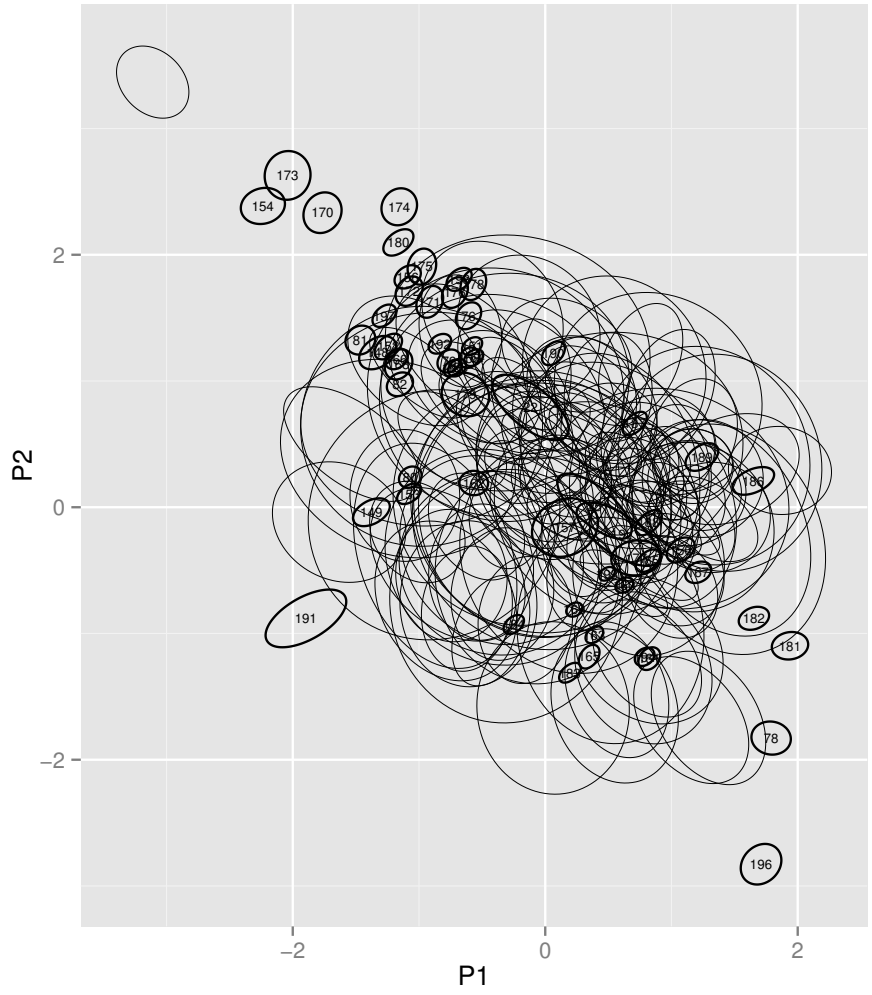


Figure 4.4: Loadings of the political questions on the factor dimensions in the model without heteroskedasticity. Ellipses are the 40% confidence ellipse for the loading estimates (roughly comparable to 1 standard deviation). Some highly uncertain estimates are displayed with thin lines to avoid completely masking questions with weak loadings.

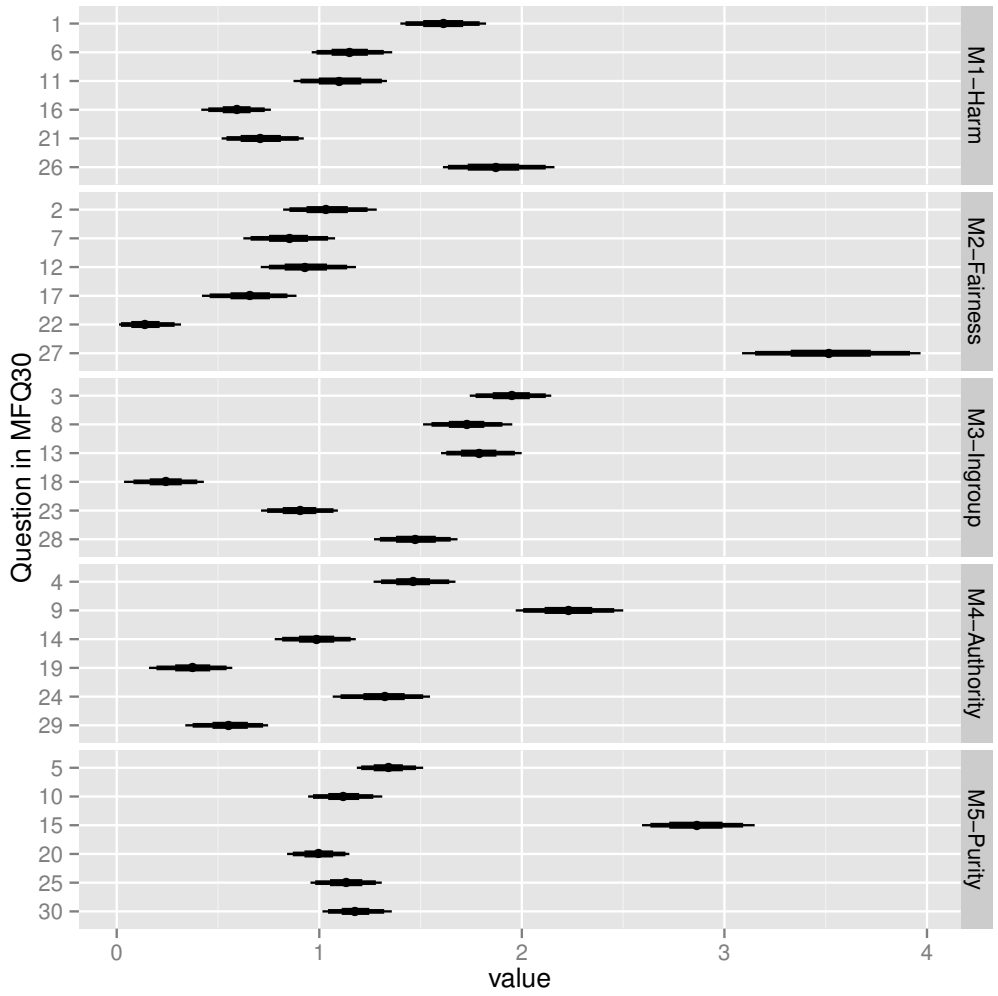


Figure 4.5: Loadings of the moral questions on their corresponding factor dimensions in the model without heteroskedasticity. Each moral question loads on only one dimension. Error bars are stepped at 60%, 80%, and 95% levels.

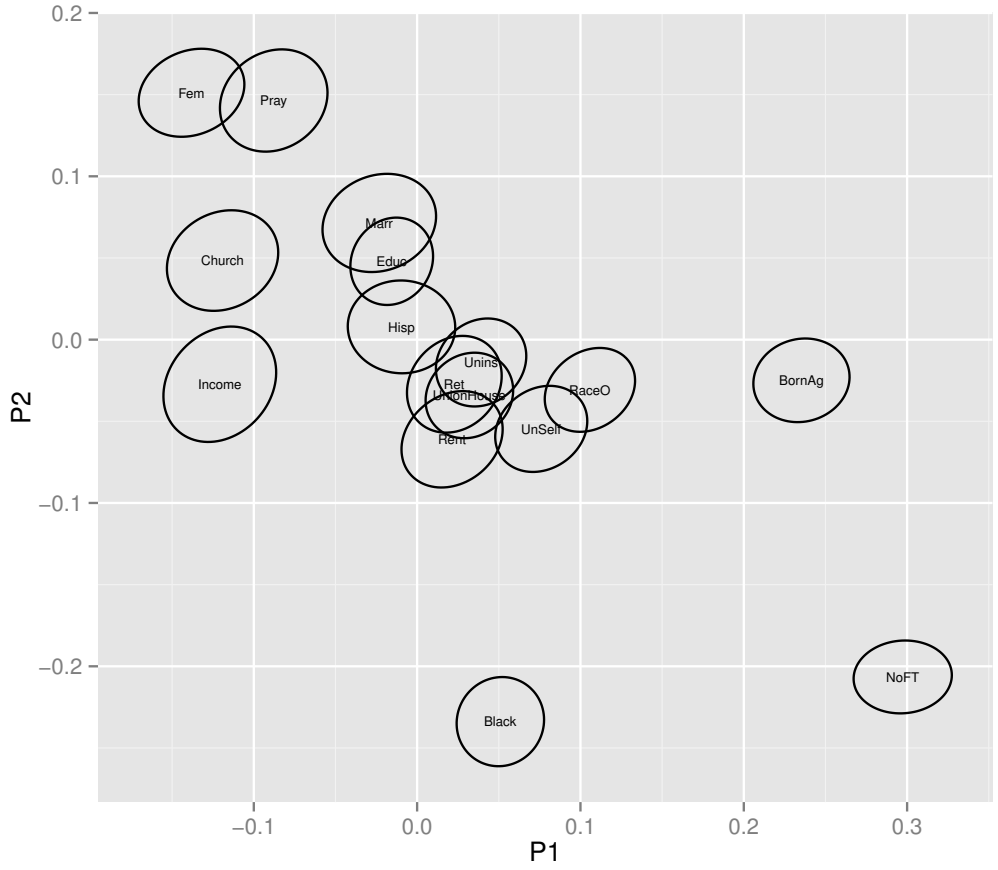


Figure 4.6: Effects of demographic parameters on political ideal points in the heteroskedastic model. Ellipses are the 40% confidence ellipse for the effect estimates (roughly comparable to 1 standard deviation).

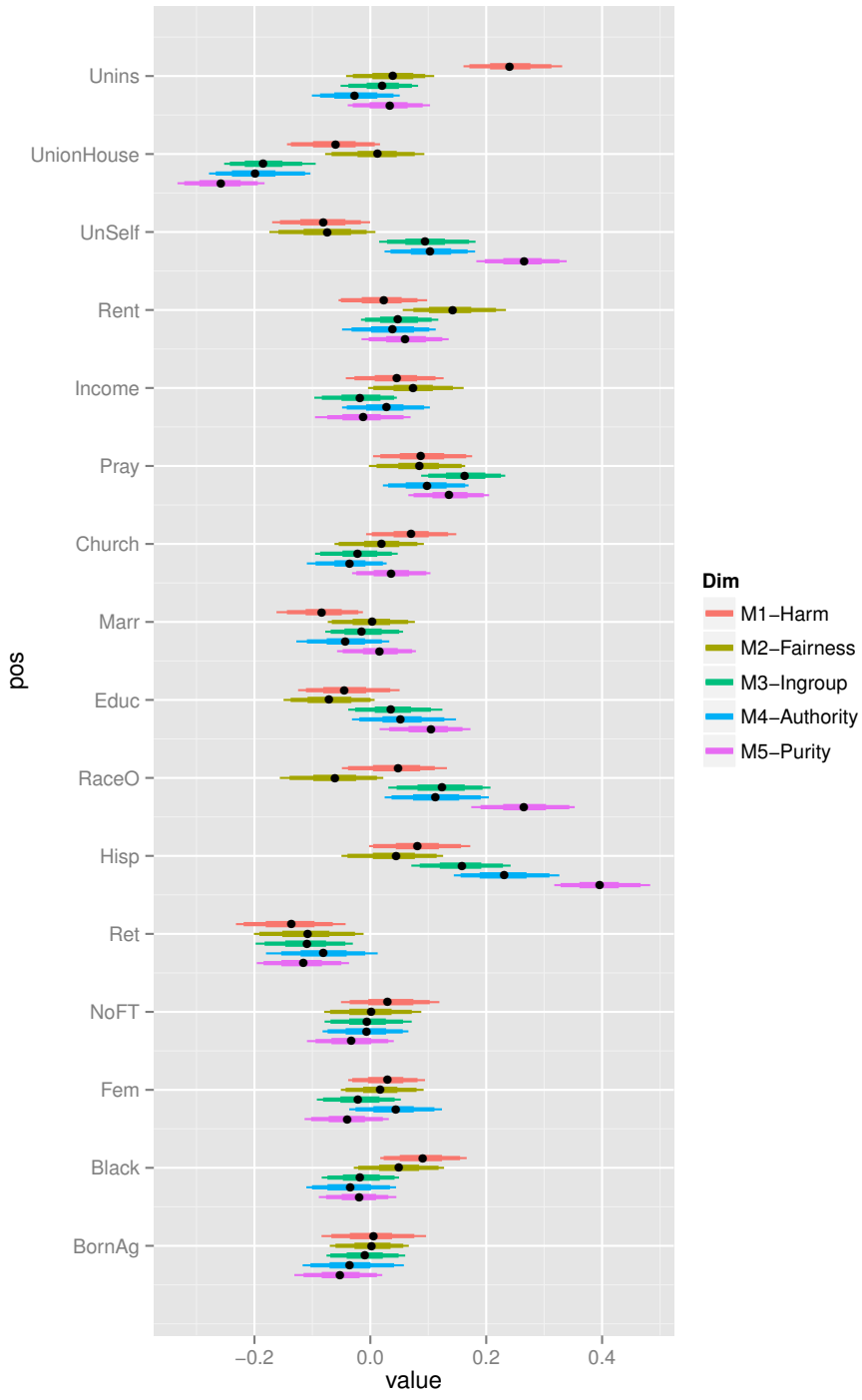


Figure 4.7: Effects of demographic parameters on moral ideal points in the heteroskedastic model. Error bars are stepped at 60%, 80%, and 95% levels.

30%, but that is primarily an artifact of the heteroskedasticity identification choice (mean $\log \sigma$ at 0) rather than a more complicated identification choice that would preserve the average scaling.

The changes are a bit easier to see in Figure 4.8, which directly compares the mean estimates of demographic effects on ideal points calculated by each model. There are definitely changes in the estimated factor structure between the two models, but they are relatively minor.

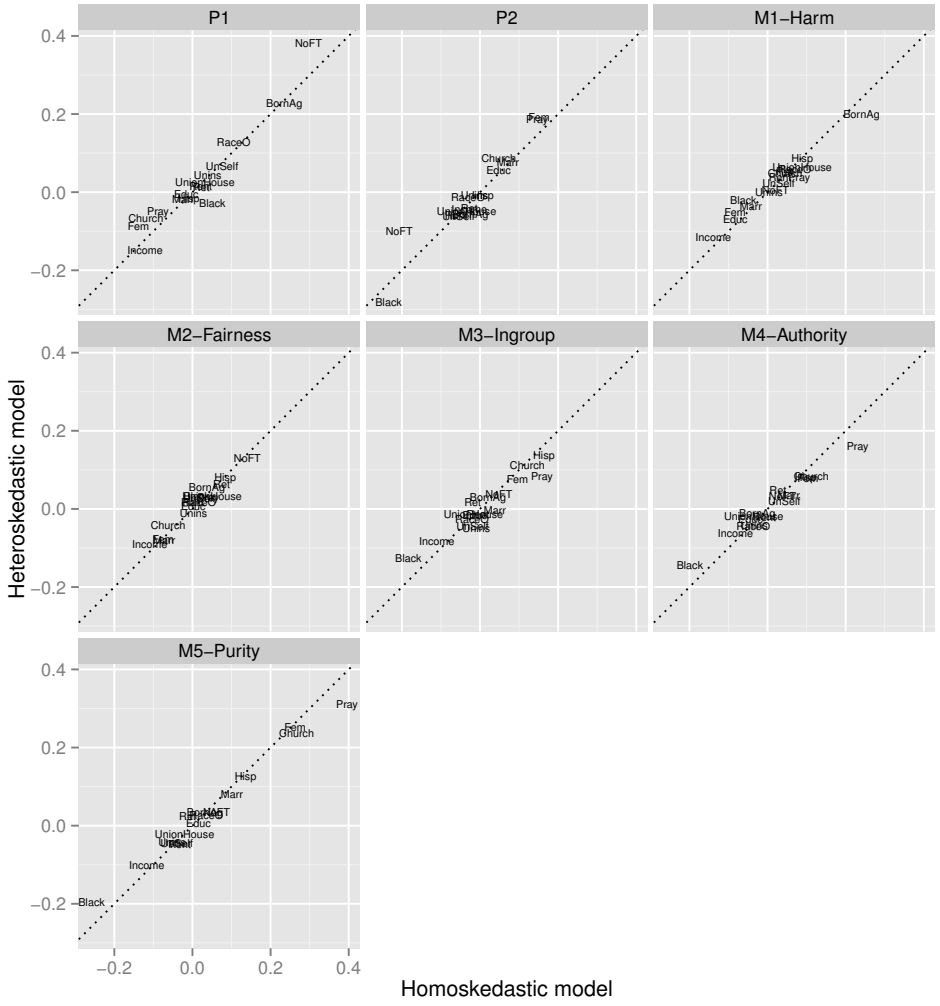


Figure 4.8: Mean posterior effects of demographic parameters on ideal points in the heteroskedastic versus the homoskedastic model

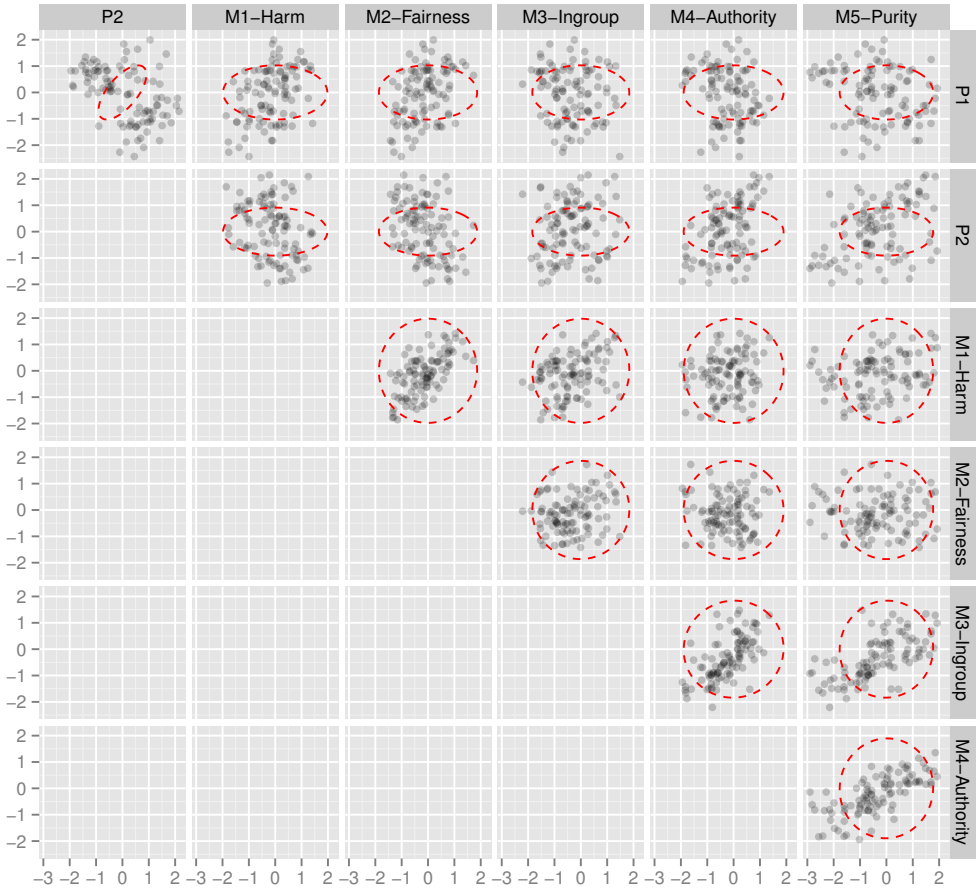


Figure 4.9: Scatterplots of a sample of the fitted ideal point in the heteroskedastic model. Red dotted ellipses denote the 95% confidence ellipse for the ideal point of a hypothetical individual at 0.

In terms of predicted ideal points, the heteroskedastic model (Figure 4.9) slightly increases the uncertainty in individual ideal point estimates compared to the homoskedastic model (Figure 4.3), but otherwise does not substantially change the qualitative properties of the ideal point estimates.

The political loadings in Figure 4.10 still show a good range of question directions, though the positions of some individual questions have moved substantially relative to the homoskedastic model (Figure 4.4), but there are too many questions to analyze in detail.

Adding heteroskedasticity to the model brings the moral factor loadings closer to their traditional scoring weights. Questions 27 and 15 are still outliers on their dimensions, but the other questions are closer together on each dimension in Figure 4.11 compared to Figure 4.5.

Finally, we move to the predictors of heteroskedasticity and the patterns of individual heteroskedasticity (Figure 4.12). The model clearly differentiates the effects of different demographics. The two largest effects on the political dimension were from NoFT, Income and Black. Looking back at Figure 4.8, we see that the effects of NoFT, Black on ideal points have more variation between the homoskedastic models than other variables. I hinted at this relationship earlier, but now the connection between movement of a parameter's ideal point effect and the parameter's effect of heteroskedasticity is readily visible.

On the moral dimensions, Pray and Black have the largest effects on moral heteroskedasticity. Looking at Figure 4.8, we see that they do shift more than the other variables, but particularly so on the Ingroup, Authority, and Purity dimensions.

Finally, we can move to the connection between political and moral heteroskedasticity. The only variable that has substantial loadings on both dimensions is Black—in that African American respondents tend to answer political questions more consistently with the 2-dimensional framework, and they tend to also answer the moral questions more consistently with the 5-factor model from the MFQ30.

There is also a connection outside of the demographics included in the model.

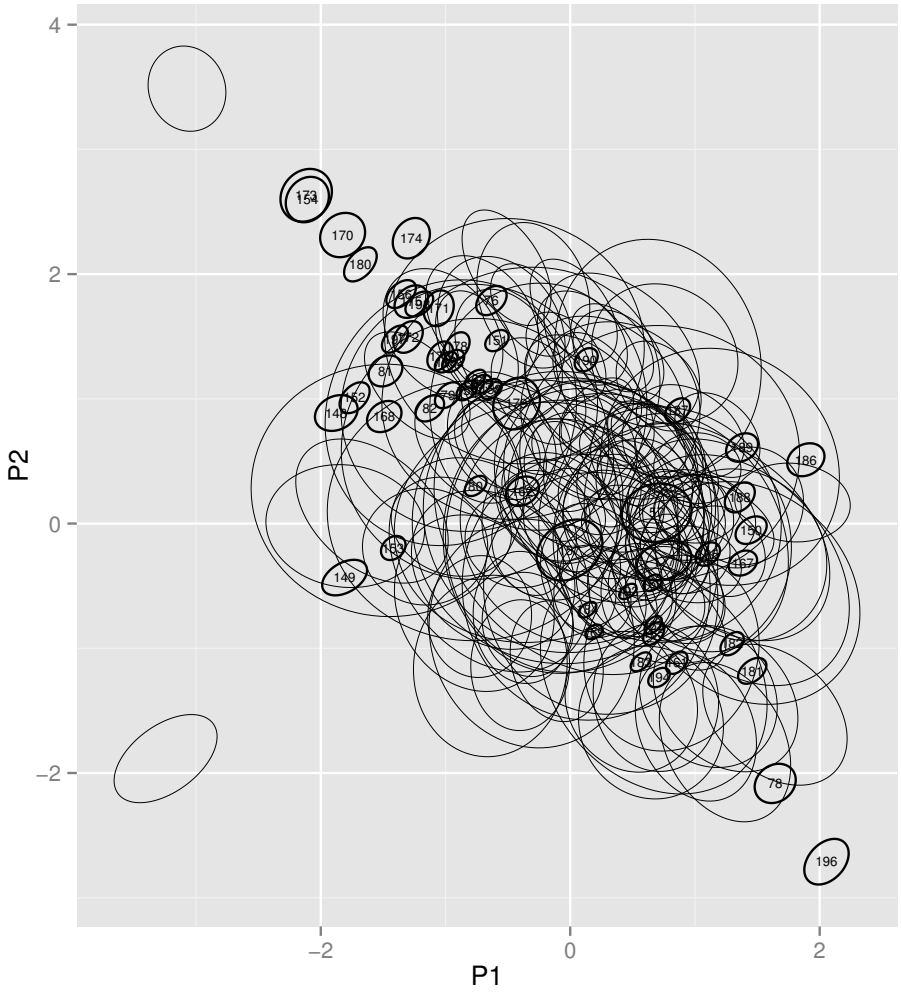


Figure 4.10: Loadings of the political questions on the factor dimensions in the heteroskedastic model. Ellipses are the 40% confidence ellipse for the loading estimates (roughly comparable to 1 standard deviation). Some highly uncertain estimates are displayed with thin lines to avoid completely masking questions with weak loadings.

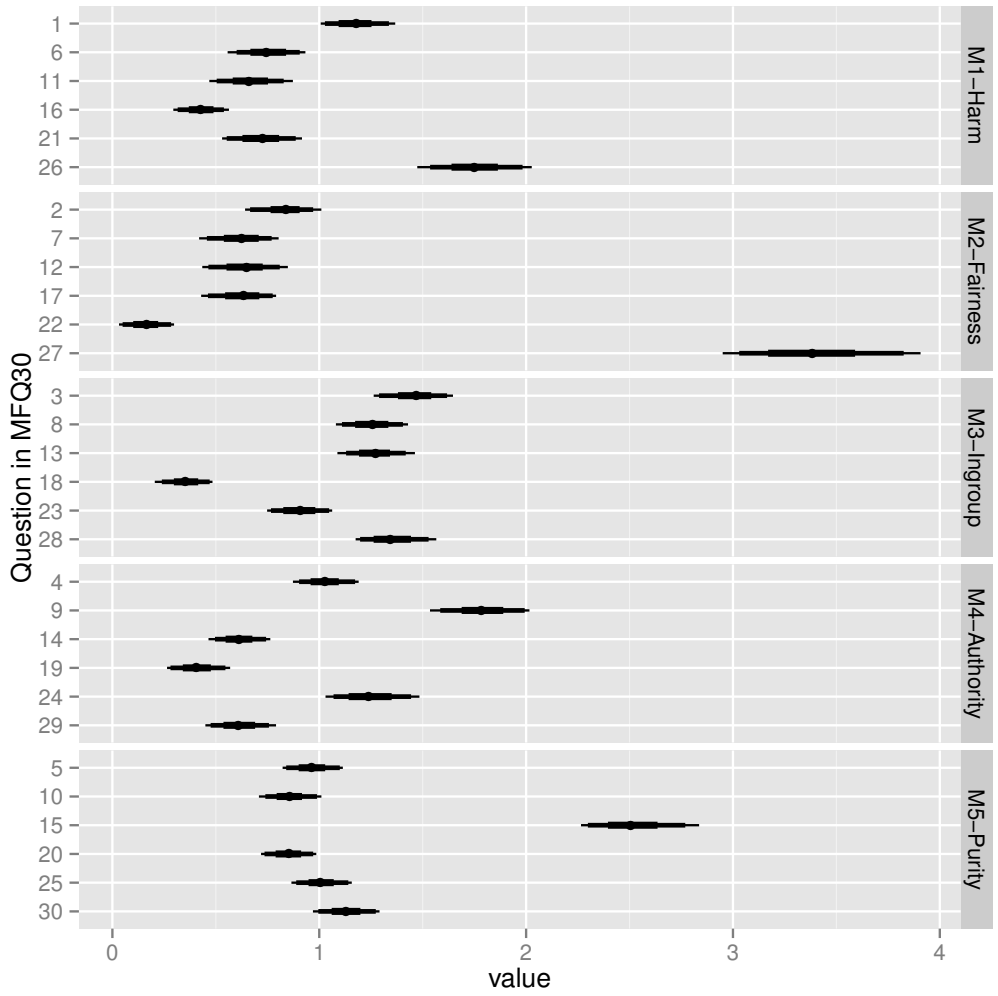


Figure 4.11: Loadings of the moral questions on their corresponding factor dimensions in the heteroskedastic model. Each moral question loads on only one dimension. Error bars are stepped at 60%, 80%, and 95% levels.

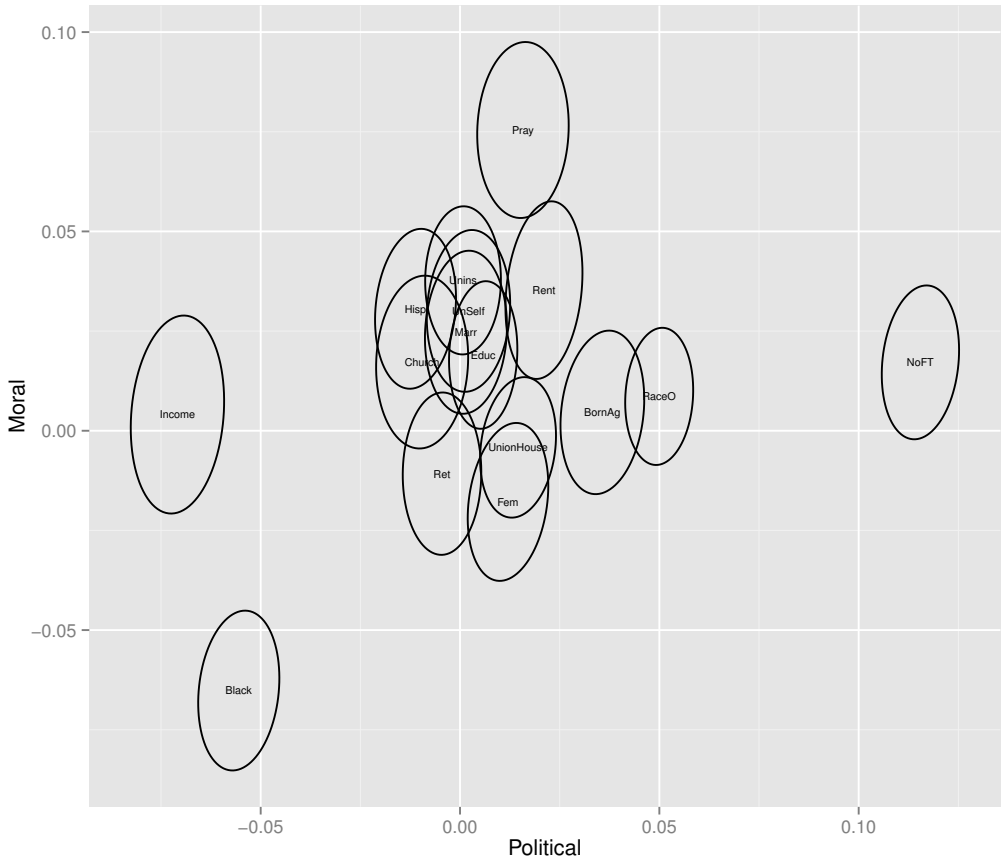


Figure 4.12: Effects of demographic variables on $\log \sigma$ of the political and moral spaces. Ellipses are the 40% confidence ellipse for the loading estimates (roughly comparable to 1 standard deviation).

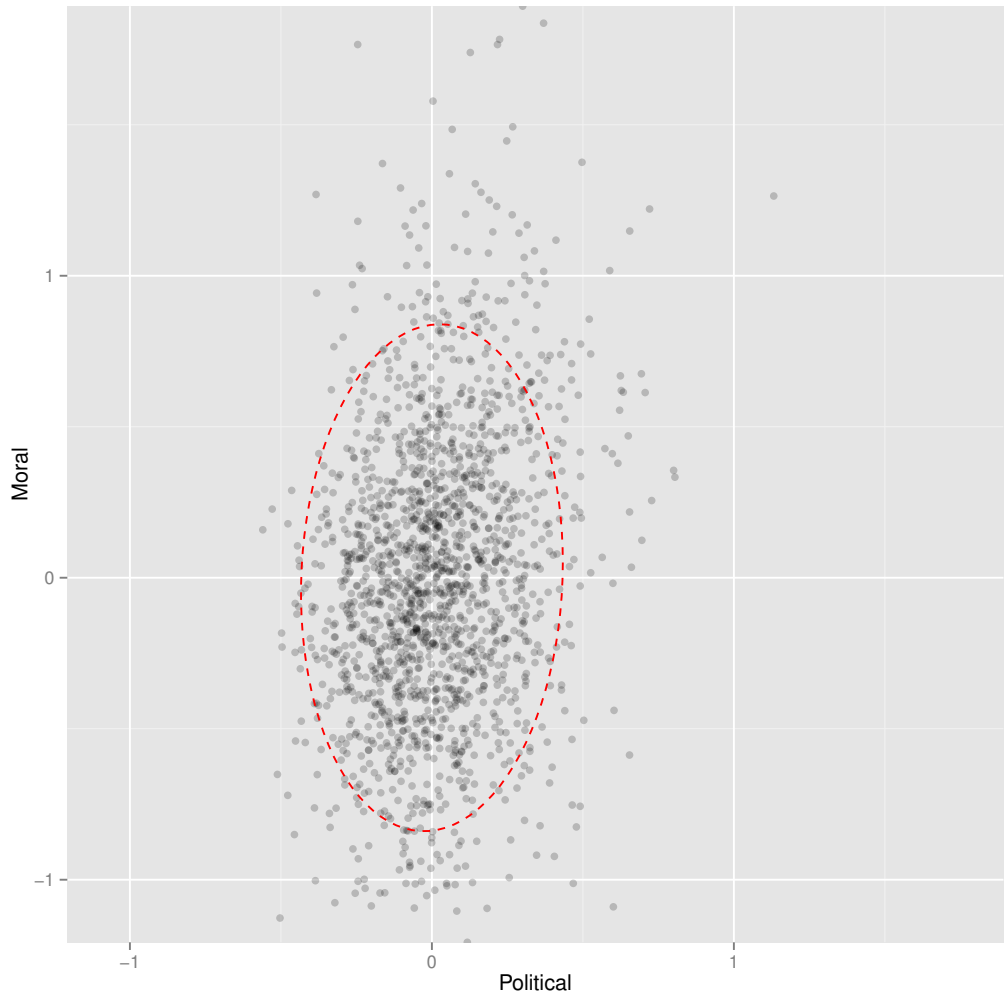


Figure 4.13: Distribution of mean posterior $\log \sigma$ for a sample of individuals. The red dotted 95% ellipse shows the uncertainty in estimates for a representative individual at the origin.

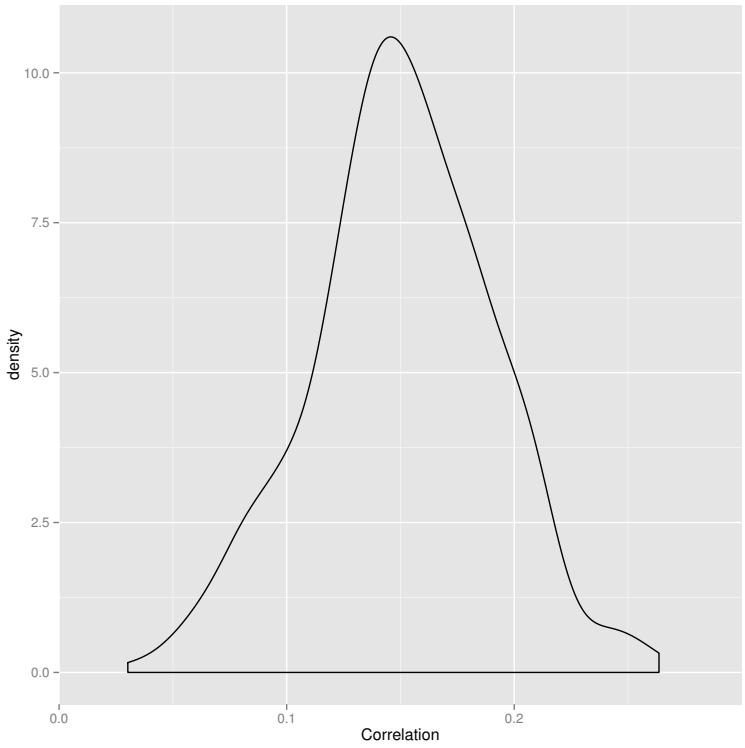


Figure 4.14: Distribution of correlation between $\log \sigma$ individual noise on the political and moral dimensions

The political and moral individual-specific noise terms on $\log \sigma$ were roughly between 0.07 and .25, so there is definitely room for more work to tease apart the missing determinants of high-level consistency in survey responses.

4.5 Conclusions

This chapter had two purposes: to understand the extent to which 'consistent' political ideology is specific to politics rather than part of a more general effect, and to construct a modeling method that allows us to jointly estimate positions in an latent space as well as consistency with that space.

On the latter, this chapter has clearly succeeded. The modeling method is tractable in real-world surveys, and it can differentiate people and groups in terms of their ideal points and levels of consistency. There is no evidence that the introduction of heteroskedasticity produces intractable posterior likelihoods or other vast departures from the standard world of ideal point modeling—the differences in the models shown here are just among specific groups that the standard modeling approaches fit poorly.

In terms of understanding the nature of political consistency, this chapter is less conclusive. There is definitely a component of political consistency that is connected to moral consistency, and this chapter shows that a connection exists in some demographic parameters and in the individual-level variation not explained by demographics, but the roots of the correlated noise are unknown.

For future applications, there are two key benefits to this model. It produces reasonably tight estimates of demographic effects on heteroskedasticity, which mean that more demographic predictors can be tested in observational work. Second, it opens the door to experimental manipulations of consistency. If random treatments applied in a survey experiment, the treatment term can be added as a hierarchical parameter to estimate a causal impact of the treatment on consistency, and that will pin down the actual causal factors underlying political consistency much more clearly than any observational study.

Finally, the key benefit of the model is that it places a hierarchical structure on political consistency and provides multiple options for linking political consistency other features of interest through the Σ^{σ} matrix, by adding hierarhical covariates to Z, or by completely replacing the construction in Equation 4.4 with some other parameterization that may be more appropriate for a specific analysis. That linkage is extremely flexible, so the model provides an straightforward path to directly relate political consistency to a wide variety of other traits or behaviors.

Chapter 5

Multidimensional salience in a model of spatial ideology

People care about different things in politics, and when they are asked a question, each individual may process it differently or call on different memories and ideas when formulating their response. In the political science literature, this heterogeneity is usually described as salience—memories, ideas, considerations, or choice features that an individual is likely to apply in their decision are considered more salient.

In the universe of spatial ideal point models of ideology, salience takes on a more specific meaning—the weights that people apply to various preference dimensions when making their choices. To the extent that 'memories and ideas' drive peoples' positions on these dimensions and the likelihood that they will use a particular dimension when making a political choice, the spatial model definition meshes reasonably nicely with the more general definition.

Formal models of voter choice and party strategy capture salience is a very compact and high-level way. Individuals have ideal points in a policy space, they are faced with possible outcomes in the policy space, and their utility of each outcome depends on the salience-weighted distance between their ideal point and the policy of interest. See Hinich and Munger (1997) for a more extensive discussion. Unfortunately, the existing techniques for spatial ideal point estimation are poorly suited to the analysis of survey data, and existing techniques for modeling

'salience' do not operate on the same latent space as salience in formal models.

In applied research on voter choices, 'salience' is almost always related to very specific ideas, considerations, or features of a model. For example, Transue (2007) emphasizes racial identity and examines how that impacts issue preferences. Gerber *et al.* (2010) emphasizes party identity in a field experiment and examines how that impacts attitudes. Outside political ideology to the related literature on latent moral preferences, Rosenblatt *et al.* (1989) demonstrate that mortality salience impacts how people punish others who violate norms.

These applications are extremely specific, and do not relate directly to high-level summaries of political ideology. The disconnect is apparent in the conclusion of Transue (2007), where the author attempts to connect the specific policy-specific findings to more general psychological and political theories. That connection is desirable, since it is much easier for us to think about a few dimensions of ideology rather than the thousands of possible policy dimensions, but so far, there has been no way to connect observable outcomes to the salience of these dimensions to individuals or groups.

Spatial ideal point estimation procedures almost never allow for individual variation in the underlying space (Clinton *et al.*, 2004; Martin and Quinn, 2002; Peress, 2009; Poole and Rosenthal, 1985), though Lauderdale (2010) is a recent exception that lets legislators vary in their consistency with the latent space as a whole, but not with specific dimensions or along diagonals. Very few estimation methods allow for demographic predictors of ideal points that are extremely important for analyzing voter behavior where we have a limited number of observations per individual and are primarily interested in group-level effects (though see Martin *et al.*, 2011, for a 1-dimensional exception).

Other researchers have looked at voter heterogeneity directly, but have taken the latent space as fixed. Rivers (1988) uses voters' self-reported ideology and partisanship and rank-ordered ratings of politicians as inputs, and examines how voters vary in how they appear to weight partisanship versus ideology. The downsides of this analysis are that the salience parameterization requires no interaction

between ideology and partisanship, and it does not produce a latent space that is at all comparable to legislative models (which were still in their infancy in 1988), so there is no way to link the results of Rivers (1988) with legislative behavior. More recently, Glasgow (2001) examined British voter choices among multiple parties given fixed and known ideal points, but allows voter heterogeneity through random coefficients on the impact of social class. As with the examples of 'salience' research described above, the heterogeneity is extremely structured and limited.

This chapter attempts to bridge the gap between the estimation of ideal points and ideological space and the estimation of salience in that ideological space. I build on constructions of ideology from the formal literature and I also build on the ideal point estimation techniques from the legislative and judicial voting literatures, to produce an intuitively interpretable and reasonably practical (though computationally intensive) method for simultaneously estimating ideal points and salience with hierarchical predictors from survey data. Unlike the existing ideal point estimation models, I allow the latent space to be 'warped' by an individual's salience matrix, and since the focus is on demographic groups rather than specific individuals, I also include hierarchical predictors of both ideal points and salience. Unlike past literature on voter heterogeneity, the model does not require an *a priori* fixed structure on the latent space and does not require that individual ideal points be known with certainty.

In the end, the procedure developed here produces group- and individual-level estimates of ideal points and salience that can drop directly into formal models like those in Hinich and Munger (1997). This allows researchers to finally test some of the models' predictions, and to the extent that the formal models hold, this model opens up new empirical avenues for understanding how parties select platforms, decide what to emphasize, choose who to target, and react to changes in voter preferences.

5.1 Building up a statistical model with salience

The aim of this chapter is to develop a link between formal models of issue salience and measureable data. The primary consideration is minimizing changes required at each end of the link—finding a construction of the statistical model that varies as little as possible from standard estimation techniques, but that also produces parameters that can drop cleanly into formal models.

There is some variation in the notation used for IRT models in political science, so to avoid confusion, Table 5.1 defines the basic IRT notation that I will use. In most multidimensional IRT models of political choices, each question has a direction vector λ and a set of cutpoints ζ, and each individual has an ideal point ϕ. The ordinal answers are given by y, and a linear latent term is given by y^*. The basic decision process follows like this[1].

$$y^* = \phi\lambda + \epsilon \quad , \quad \epsilon \sim \text{Normal}(0, 1) \tag{5.1}$$

or equivalently

$$y^* = \left(\phi + \frac{\epsilon}{\|\lambda\|}\right)\lambda \quad , \quad \epsilon \sim \text{MVN}(0, I) \tag{5.2}$$

and then

$$y = 1 + \sum_j \{y^* < \zeta_j\} \tag{5.3}$$

The equivalence of the 5.1 and 5.2 is shown in more detail in Appendix B.1. The intuition behind this model is that a low-dimensional ideological space summarizes basic patterns of policy preferences in the population. A particular individual has an ideal 'summary' policy bundle represented by a point in this low-dimensional space. Their answers to particular question are based on their summary policy plus some question-specific variation that is not captured by the summary policy.

Note that in this construction, everyone employs the same 'summary' space when considering policies, and question-specific variation is highly constrained—independent and homoskedastic—and those constraints may not reasonably cap-

[1]There are several other decision processes that produce observationally equivalent outcomes, but this notation will simplify later discussions.

Table 5.1: Model notation

s	index of the 'subject' or person
q	index of a specific question or choice
ϕ_s	s's ideal point in the latent ideological space
y^*_{sq}	s's ideal response on the single-dimensional choice line for question q
y_{sq}	s's actual response or choice on question q, indexed from 1
J_q	the number of possible responses to question q minus 1, which is the number of cutpoints in an ordinal model
ζ_{qj}	the position of the jth cutpoint on the single-dimensional choice line for question q
λ_q	a nonzero vector that maps the ideal points onto question dimensions
A_s	the salience weighting matrix for subject s
ϵ_{sq}	the question-specific random noise used by subject s when answering question q

ture political thought.

5.1.1 Linking IRT and formal models

The most basic formal models that employ spatial voter preferences treat ideal points as common knowledge, use a common policy space shared by everyone, and treat the voter choices as perfect optimizations based on euclidean distance in the policy space—meaning that voters have discrete or intentionally mixed strategies, but they do not have any random noise added in to their decisions.

Policy space

Importantly, the 'common policy space' in the formal models is pre-defined. There is no uncertainty in the spatial parameters. In the statistical domain, this corresponds to confirmatory factor analyses that define a few policy dimensions, assign questions to each dimension, and then try to fit ideal points and other parameters into the pre-defined space. However, the assignment of policy dimensions is always somewhat ad hoc (or it would be exploratory factor analysis) and subject to biases on the part of the analyst.

Because there is generally to little information available to pre-assign the policy dimensions to questions, the legislative and judicial voting literatures have tended towards exploratory models that do not constrain the policy space more than needed for statistical identification. I view this as a wise choice, and so the linkage between formal and statistical models I develop here is based more on exploratory and less on confirmatory factor analyses.

Hyperplanes versus outcome positions

The next hurdle to be addressed is the nature of voter choices. Formal models of voter choice usually define options as outcome positions in policy space, and voters want to obtain the policy outcome closest to their ideal point.

Formally, those models define outcomes ν_{qj} that fall along a line defined by vectors α_q^ν and λ_q with $\alpha_q^\nu \cdot \lambda_q = 0$ and positions on that line defined by scalars

γ_{qj}^{ν}.

$$j \in \{1, \ldots, J_q + 1\} \tag{5.4}$$

$$\nu_{qj} = \alpha_q^{\nu} + \lambda_q \gamma_{qj}^{\nu} \tag{5.5}$$

$$y = \underset{j \in \{1,\ldots,J_q+1\}}{\arg\min} (x_{sq} - \nu_{qj})^{\mathsf{T}}(x_s - \nu_{qj}) \tag{5.6}$$

IRT models, on the other hand, use hyperplanes to split the space into groups with different question-specific preferences $x_{sq} = \phi_s + \frac{\epsilon_{sq}}{\|\lambda_s\|}$. These formulations can be observationally equivalent for binary and ordinal choices with up to 4 options—meaning given a set of hyperplanes, we can construct policy outcomes such that the voters between two hyperplanes are closer to the corresponding outcome position than to any other outcome position.

When the number of choices is 5 or higher—meaning there are 4 or more hyperplanes—it is possible to have hyperplanes that split the cloud of voter preferences in ways that are impossible to obtain from outcome positions alone.[2] We can, however, represent a model that uses ordered outcome positions with a hyperplane construction by constructing hyperplanes perpendicular to the line of outcome positions that intersect the line at the midpoints between neighboring outcomes.

Thus, any data that can be represented by outcome positions can be represented equally well, if not better, by hyperplanes, and so I will use a hyperplane representation in the modeling. To the extent that the model's fitted parameters can be represented by outcome positions, if we want to analyze questions in terms of their outcome positions, it is possible to translate fitted hyperplanes into a set of policy outcomes for each question. So when the policy outcome construction is reasonable in light of observed data, we can calculate the outcome positions, but when the data is inconsistent with outcome points, the model is not constrained.

[2]Example: with unidimensional space, put cutpoints at -1.1, -1, 1, and 1.1. There is no point ζ such that $\|\zeta - (-1)\| < 0.1$ and $\|\zeta - 1\| < 0.1$, so there is no way that distance minimization can make the central region wide enough relative to the width of the neighboring regions

Before moving on, I should note that with choice-specific utility noise, it is possible to represent any set of hyperplanes, but that nearly doubles the degrees of freedom per choice and greatly complicates identification. Such a setup might be useful in ex-post conversion from hyperplanes to ideal points, but it is not helpful within the model itself.

Noise parameterization

In formal models, random variation in voter choices is either ignored in the model, or it usually enters as outcome-specific biases that behave like the noise term in multinomial logit. In binary and ordinal IRT models, noise enters the voter choice as a bias on the question-specific latent term (y^* in this chapter's notation). In the MNL-style formal models, if we hold all other parameters fixed, increasing the amount of noise in the model will shift all outcome probabilities monotonically towards a uniform distribution. In ordinal probit models used in IRT, extremely high levels of noise will shift the outcome probabilities towards the extreme values and away from middle options, but at moderate levels of noise, the impact on outcome probabilities is non-monotonic. As noise variance increases, nearby interior options can rise in probability at first, but then fall as the variance increases enough to shift all the probability towards the extreme options.

In the models developed here, I base the noise parameterization on a traditional ordinal IRT construction and only modify when necessary to introduce salience into the model. This is the standard method used in the limited literature on political ideal point estimation with ordinal data, but there is no principled reason to assume that the IRT construction of noise is superior to MNL-style biases. Optimal noise parameterization is an empirical question that merits further research, but it is outside the scope of this chapter, and so I stick to the standard, though possibly imperfect, parameterization.

5.1.2 Adding in Salience

In formal models that use outcome policy positions, salience enters in the distance calculation. For a positive definite salience matrix A_s, ideal point ϕ_s, and policy outcomes $\nu_1, \ldots, \nu_{J_q}$, we just make a slight tweak to Equation 5.6:

$$\begin{aligned} \nu_{qj} &= \alpha_q^\nu + \lambda_q \gamma_{qj}^\nu \\ y &= \underset{j \in \{1,\ldots,J_q+1\}}{\arg\min} (x_s - \nu_{qj})^\mathsf{T} A_s (x_s - \nu_{qj}) \end{aligned} \tag{5.7}$$

As we have discussed, when $A = I$, the parameterization in terms of outcome policies ν is a constrained version of a hyperplane model. We do not want to force those constraints just so that we can model salience, so we must instead find a parameterization of salience that works in the hyperplane representation. To do this, I will first parameterize 5.7 in terms of hyperplanes, and then see what adjustments (if any) need to be made so that the hyperplane version can be unconstrained.

We will look at the simplest case for optimization. We have two outcomes points ν_1 and ν_2, and we want to find the hyperplane that splits the space of x_s according to their optimal outcome. For simplicity, I will drop some constant subscripts in this derivation. The values on the hyperplane z satisfy

$$(z - \nu_1)^\mathsf{T} A (z - \nu_1) = (z - \nu_2)^\mathsf{T} A (z - \nu_2), \tag{5.8}$$

and since A is positive definite, we can find a square root matrix with the same eigenvectors using an eigendecomposition:

$$A = Q \Lambda Q^\mathsf{T} \tag{5.9}$$

$$B = Q \Lambda^{1/2} Q^\mathsf{T} \tag{5.10}$$

where Q is an orthonormal matrix, Λ is a diagonal matrix of positive eigenvalues,

and $\Lambda^{1/2}$ is a diagonal matrix of the square roots of the eigenvalues.[3] Then B is positive definite and has the property that

$$A = B^{\mathsf{T}}B = BB^{\mathsf{T}}. \tag{5.11}$$

We substitute $B^{\mathsf{T}}B$ for A in equation 5.8 to get

$$(z-\nu_1)^{\mathsf{T}}B^{\mathsf{T}}B(z-\nu_1) = (z-\nu_2)^{\mathsf{T}}B^{\mathsf{T}}B(z-\nu_2) \tag{5.12}$$

$$(Bz-B\nu_1)^{\mathsf{T}}(Bz-B\nu_1) = (Bz-B\nu_2)^{\mathsf{T}}(Bz-B\nu_2) \tag{5.13}$$

which is straightforward to solve in terms of Bz:

$$Bz = \frac{B\nu_1 + B\nu_2}{2} + C \tag{5.14}$$

where

$$C \cdot (B\nu_1 - B\nu_2) = 0. \tag{5.15}$$

Since $\nu_1 - \nu_2$ is proportional to λ as defined in Equation 5.5, $(B\nu_1 - B\nu_2)$ is proportional to $B\lambda$, and so the hyperplanes must be of the form

$$z^{\mathsf{T}}B\lambda = \text{constant} \tag{5.16}$$

To find the constant, all we have to do is calculate the value under the z for which

[3] A more common 'matrix square root' is the Cholesky decomposition, but the Cholesky decomposition is not positive definite, nor does it preserve eigenvalues, so it can rotate and skew the space rather than simply scaling it along set dimensions. Also, the computations involved in fitting the model primarily calculate A from B rather than the reverse, so this particular eigendecomposition is not actually costly in terms of computation.

$C = 0$ in Equation 5.14, which I will denote by z^0.

$$z^0 = B^{-1}\left(\frac{B\nu_1 + B\nu_2}{2}\right) \tag{5.17}$$

$$= \frac{\nu_1 + \nu_2}{2} \tag{5.18}$$

and we can then punch in z^0 to solve for the constant in 5.16

$$(z^0)^\mathsf{T} B\lambda = \left(\frac{\nu_1 + \nu_2}{2}\right)^\mathsf{T} B\lambda \tag{5.19}$$

$$= \left(\frac{\alpha + \lambda\gamma_1 + \alpha + \lambda\gamma_2}{2}\right)^\mathsf{T} B\lambda \tag{5.20}$$

$$= \left(\alpha + \lambda\frac{\gamma_1 + \gamma_2}{2}\right)^\mathsf{T} B\lambda \tag{5.21}$$

Note that the α term cannot be reduced away; while $\alpha^\mathsf{T}\lambda = 0$ by definition, it is not generally true that $\alpha^\mathsf{T} B\lambda = 0$. And since the B terms vary between individuals, the 'constant' on the right hand side is not the same for each individual.

This highlights the fact that nonuniform salience removes the 'perpendicular shift invariance' property described in Appendix B.2. That means that aside from any parameters that determine A_s itself, we also double the number of parameters associated with each question. Now rather than just having a direction, each question has a direction and an origin.

But we can simplify this slightly. $\left(\alpha + \lambda\frac{\gamma_1+\gamma_2}{2}\right)$ can just be written as $(\alpha + \lambda * \eta_j)$ with scalar η for each cutpoint (which is now a point in latent space rather than a point on the real line). Much like the outcome points, these cutpoints fall on a line in the latent space, but the cutpoint construction again offers more flexibility in the positioning of slicing hyperplanes just as they did in the version without salience.

The final task is to translate back from hyperplanes into a model structure like equations 5.2 and 5.3.

$$y^*_{sq} = \left(\phi_s + \frac{\epsilon_{sq}}{\|\lambda\|}\right) B_s \lambda_q, \qquad \epsilon_{sq} \sim \text{MVN}(0, I) \tag{5.22}$$

and then
$$y_{sq} = 1 + \sum_j \left\{ y^*_{sq} < (\alpha_q + \lambda_q * \eta_{qj})^\mathsf{T} B_s \lambda_q \right\} \tag{5.23}$$

5.1.3 Noise parameterization

Most IRT models assume homoskedastic latent errors, but the salience construction above introduces structured heteroskedasticity that needs to be examined before we move forward. Equation 5.22 can be rewritten as

$$y^*_{sq} = \phi_s B_s \lambda_q + \frac{\epsilon_{sq} B_s \lambda_q}{\|\lambda\|}, \qquad \epsilon_{sq} \sim \text{MVN}(0, I) \tag{5.24}$$

and then the noise term is distributed as

$$\frac{\epsilon_{sq} B_s \lambda_q}{\|\lambda\|} \sim \text{Normal}\left(0, \frac{\lambda^\mathsf{T} B_s^\mathsf{T} B_s \lambda}{\|\lambda\|^2}\right) \tag{5.25}$$

$$\sim \text{Normal}\left(0, \frac{\lambda^\mathsf{T} A_s \lambda}{\|\lambda\|^2}\right) \tag{5.26}$$

and the cutpoint transformation can be expanded similarly

$$(\alpha_q + \lambda_q * \eta_{qj})^\mathsf{T} B_s \lambda_q = \underbrace{\alpha_q B_s \lambda_q}_{\text{shift}} + \underbrace{\lambda_q^\mathsf{T} B \lambda_q}_{\text{scaling}} \eta_{qj} \tag{5.27}$$

We want to compare the scaling term in Equation 5.27 with the variance in 5.26, both of which are scalars. In the simplest case, where A_s is the identity matrix, the variance in 5.26 equals 1, and the scaling term in 5.27 is equal to $\|\lambda\|^2$. As we move away from $A_s = I$, the variance of the noise in 5.26 increases proportionally with the A_s-weighted length of λ_q. The scale term increases proportionally to the B_s-weighted length of λ_q. Even though the eigenvectors of A_s and B_s are the same by construction, there is not a simple linear relationship between $\lambda_q^\mathsf{T} A \lambda_q$ and $\lambda_q^\mathsf{T} B \lambda_q$.

To illustrate this, examine a simple 2-dimensional case with

$$A = \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} \sqrt{s} & 0 \\ 0 & 1 \end{pmatrix}$$

$$\lambda_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \lambda_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \lambda_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

where s is an extra term to illustrate what happens as A moves away from I. As s increases, the scaling terms $\lambda_q^\mathsf{T} A \lambda_q$ and $\lambda_q^\mathsf{T} B \lambda_q$ vary as shown in Figure 5.1.

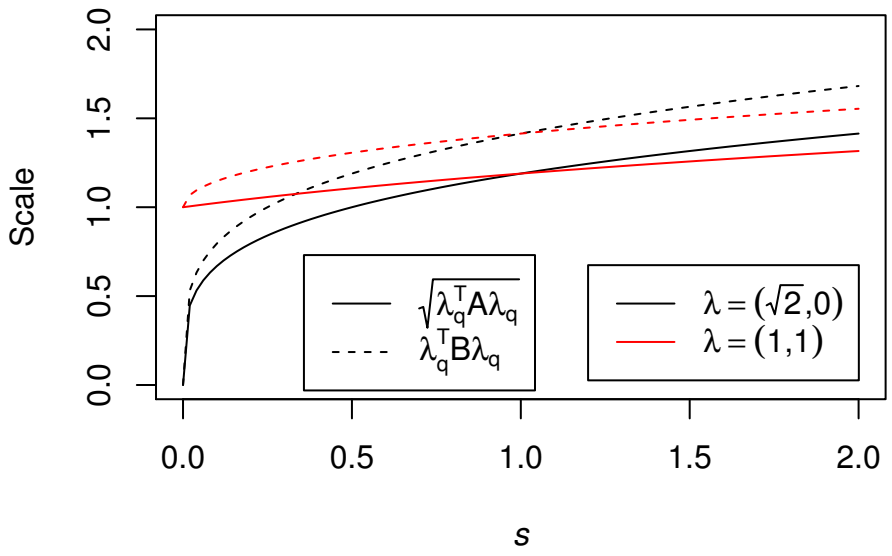


Figure 5.1: Changes in error standard deviation and cutpoint scaling terms as the salience matrix moves away from the identity matrix

To a very rough approximation, increases in salience in the direction of or perpendicular to λ_q do not have a large impact on the noise in the question-specific ordinal model. When λ is in between parallel and perpendicular to the salience changes, movement away from I acts like a very slight increase in the noise of the question-specific ordinal model, but the effect is only apparent at rather extreme salience weights. That does not mean that salience has no impact on the model, just that its primary effect will often be through the shift term in Equation 5.27.

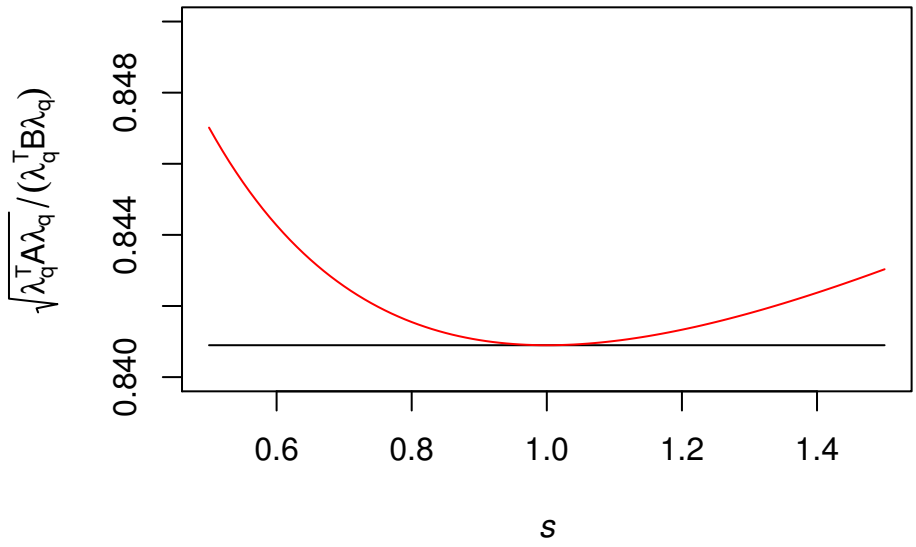


Figure 5.2: Change in ratio of error standard deviation to cutpoint scaling terms as the salience matrix moves away from the identity matrix

If additional flexibility is desired in the model, a natural place to add it is in this noise parameterization term—while the parameterization in this chapter does not substantially vary the ordinal model's noise term, we might want it to. For example, if we think that people near extremes on particular dimensions would collect more information relating policies to those dimensions, then they might have very clear opinions on policy questions that lie on their 'favorite' dimension and would have lower noise in their question-specific ordinal model. That sort of extension is left for future work, however, and this chapter focuses on developing the basic model.

5.1.4 Hierarchical modeling of salience matrices

With a relatively small number of noisy observations per individual, non-hierarchical estimates of ideal points can be quite noisy, and individual-level estimates of salience are even more disconnected from actual observable data. Furthermore, since we are more interested in group-level and population-level variation than in

the positions of specific survey repspondents, we need a hierarchical structure to determine what patterns of salience exist in the population as a whole.

Unfortunately, there are no directly applicable models for us to use here. Other fields that use hierarchical models of positive definite matrices usually focus on shrinking submatrices towards a shared 'parent' matrix or towards a particular global structure. Here our problem is different—we want to use a set of continuous variables to predict an individual's salience matrix, but we do not have a nested structure to our data or other 'parent' matrix that might apply. So, we need to design our own modeling technique for positive definite matrices.

The first place to start is in determining what properties we need from a hierarchical model of salience matrices:

- Salience matrices should be positive definite
- Reordering demographic predictors should have no impact on the resulting salience matrices
- Removing a demographic predictor from the model should have the same impact as having a 'zero' parameter for that predictor, where 'zero' remains to be defined
- Demographic predictors should not have to impact the same set of eigenvectors, etc.
- Demographic predictors should be able to increase or decrease salience in any direction

Order invariance is a particularly difficult property to obtain since very few operations on matrices are commutative. Matrix addition does commute, but the sum of positive definite salience matrices has strictly greater salience on all dimensions than any of its components, and so simple addition of matrices will not let demographic predictors decrease salience on a specific dimension. Yet order invariance is one of the most important elements of this model. Without it, we have no simple way of interpreting the results, since the interpretation of each

demographic's effects depends on the effects of all the other demographics in the model and the sequence in which they are included in the model. It also greatly complicates the construction of priors since priors that are reasonable when only one demographic is included in the model may have very different meanings when other demographics are added.

For example, one naive approach to combining positive definite matrices is to simply multiply them. Here, that means we would construct positive definite matrices for the effect of each demographic on each individual and multiply all of an individual's matrices to obtain their final salience matrix. Unfortunately, if the eigenvectors of all of these matrices are not identical, the second and later demographic variables are effectively operating on a rotated version of the ideal point space rather than simply a scaled one. At the extreme, if the demographic matrices have multiplied together to rotate the space significantly, a demographic that appears to increase salience on one axis—for example, a diagonal matrix with a value of 2 at $(1, 1)$ and 1 on the rest—may actually be increasing the salience of a completely different axis.

The order invariance, eigenvector variation, increase/decrease, and 'zero' parameter properties are all simple to satisfy if we can just operate on the space of real symmetric matrices rather than positive definite ones. If we can aggregate the demographic effects in a symmetric matrix space and transform that into a space of positive definite matrices, then we can satisfy all the desired properties fairly easily.

Conveniently for us, the eigendecomposition-based matrix exponential gives a straightforward transformation from real symmetric matrices to positive definite ones. The transforms below give an straightforward and almost one-to-one mapping between the space of real square matrices and positive definite matrices.[4] B is a positive definite matrix, and E is its counterpart in the space of symmetric

[4]The mapping is one-to-one whenever B has unique eigenvalues, or when the computational eigendecomposition method gives the same eigenvectors for $Q\Lambda Q^{-1}$ as for $Q\exp(\Lambda)Q^{-1}$ where Q is orthonormal and Λ is a diagonal matrix.

matrices.

$$A = Q \exp(\Lambda) Q^{\mathsf{T}}$$
$$E = Q \Lambda Q^{\mathsf{T}} \tag{5.28}$$

where Λ is a diagonal matrix and $[\exp(\Lambda)]_{ij} = \exp(\Lambda_{ij})$.

This gives a natural way to aggregate the effects of different demographic variables. Each individual's E can simply be a sum of demographic matrices weighted by the individual's demographic parameters—weighted sums of matrices have a natural 'zero' term (the zero matrix), they are commutative, and they do not require that the component matrices share eigenvectors.

So, given a length-v vector of demographic covariates Z_s, we can use the exponential transformation to construct their salience matrix as follows. First we draw their individual-specific noise according to some distribution over symmetric matrices

$$\tilde{E}_s \sim F \tag{5.29}$$

then we calculate a weighted sum of the $k \times k$ demographic parameter matrices plus the individual-specific noise

$$E_s = \tilde{E}_s + \sum_v Z_{sv} \cdot E_s^v \tag{5.30}$$

take its eigendecomposition so that we can construct a square root matrix with the same eigenvectors

$$E_s = Q_s \Lambda_s Q_s^{\mathsf{T}} \tag{5.31}$$

and finally calculate the salience matrix and its square root

$$A_s = Q_s \exp(\Lambda_s) Q_s^{\mathsf{T}} \tag{5.32}$$

$$B_s = Q_s \exp\left(\tfrac{1}{2}\Lambda_s\right) Q_s^{\mathsf{T}} \tag{5.33}$$

5.1.5 Identification of salience matrices

Once the scale, rotation, and reflection of the latent space are identified by any of numerous standard methods (see Rivers (2003) for a summary of options), all that remains for identification is to ensure that the scale of salience in each dimension is standardized.

In particular, multiplication of all A_s matrices by a positive definite matrix can be counteracted by appropriate shifts in that λ and η terms, so we need to pin down its scale. One way to do this is to pin down the mean of all the E_s terms, which is easy to do by centering the matrix of hierarchical predictors Z and ensuring that Z does not include a column of constants. To aid in interpretation, it is also helpful to scale the columns of Z to unit variance, though this is not absolutely necessary.

5.2 Final model specification

I have discussed a variety of modeling options, but for clarity, I will write everything down in one place here. The model uses only two latent dimensions of ideology.

$$y^*_{sq} = \phi_s B_s \lambda_q + \epsilon_{sq}, \qquad \epsilon_{sq} \sim \text{Logistic}(0, 1) \tag{5.34}$$

$$y_{sq} = 1 + \sum_j \left\{ y^*_{sq} < (\alpha_q + \lambda_q * \eta_{qj})^{\mathsf{T}} B_s \lambda_q \right\} \tag{5.35}$$

$$\phi_s \sim \text{MVN}(Z\beta^{\phi}, I) \tag{5.36}$$

$$B_s = Q_s \exp(\Lambda_s) Q_s^{\mathsf{T}} \tag{5.37}$$

$$Q_s \Lambda_s Q_s^{\mathsf{T}} = E_s \tag{5.38}$$

with demographic predictors on the E_s terms

$$E_s = \sum_v (Z_{sv} D^v) + \tilde{E}_s, \tag{5.39}$$

with the following weakly informative priors. Note that $\tilde{E}_s$ and D^v are constrained to be symmetric, so the priors below apply to only the upper triangle, and the subdiagonal terms are constrained to the transpose of the superdiagonal terms.

$$(\tilde{E}_s)_{ij} \sim \text{Normal}(0, \sigma^h) \tag{5.40}$$

$$D^v_{ij} \sim \text{Normal}(0, \sigma^d) \tag{5.41}$$

$$\lambda_q \sim \text{MVN}(0, \sigma^\lambda I) \tag{5.42}$$

$$\alpha_q \sim \text{MVN}(0, \sigma^\alpha I), \tag{5.43}$$

The main alteration from the above discussion is that an ordinal logit model is used for the final stage rather than ordinal probit. This is done to take advantage of efficient ordinal logit code in Stan that does not have an ordinal probit counterpart, but ordinal logit and probit are rarely very different in practice, so the change is unlikely to be substantively important.

The normal priors over each term in the matrices $\tilde{E}_s$ and D^v are admittedly somewhat ad hoc, but there is no literature on the distributions of logs of covariance matrices. In principle, I could have put a Wishart prior on $\exp(\tilde{E}_s)$ and $\exp(D^v)$, but that would have added an additional $n + k$ eigendecompositions to the likelihood calculation, and that would significantly slow an already computationally demanding model.[5]

For $\sigma = 1$, and 2 dimensions, the priors on $\tilde{E}_s$ and D^v produces matrices that,

[5]Another option would be to separate the priors over the correlation and marginal variance components of the covariance matrices, which could allow the scale of the demographic's effects on salience to vary more while keeping the correlation components constrained. That offers more flexibility but unfortunately still requires an additional layer of $n + k$ eigendecompositions to calculate the matrix logs needed by the likelihood calculation.

when exponentiated, qualitatively approximate an Inverse Wishart distribution with an identity scale matrix and 3 degrees of freedom, which is a fairly common choice of prior over positive definite matrices. Smaller values of σ shrink the resulting $\exp(\tilde{E}_s)$ and $\exp(D^v)$ towards the identity matrix.

5.3 Application

I apply this model to a subset of the common content from the 2010 Cooperative Congressional Election Survey (CCES). The CCES is an internet survey composed of a number of 'team' modules from various universities and groups, plus a core set of common questions asked to all respondents. The survey is administered by Polimetrix on a representative sample of US voters. The full common content file has over 50k respondents, but I use a sample of 1000 respondents from California. I do this for two reasons: the full sample is too large for efficient computation (the CA sample takes about a day to run, even with Hamiltonian Monte Carlo); and so long as the data must be subsampled, it is helpful for interpretation to have a relatively coherent political environment across the sample. California has distinct regional politics, but there is much less variation than across the United States as a whole.

The model is estimated in Stan (Stan Development Team, 2013a,b) from R 2.15.3 (R Core Team, 2012). Stan samples from the posterior using Hamiltonian Monte Carlo, which speeds the estimation of this particular model by an order of magnitude, and it includes functionality for calculating the eigendecompositions used in this model.

5.4 Results

I run two models—one without variable salience, and a second with salience included in the model. The first model gives a baseline for comparing how introducing salience changes the demographic effects on ideal points and the factor loadings of survey questions.

5.4.1 Model without salience

For the first model, I alter the specification from Section 5.2 to fix all the B_s terms to the identity matrix. This turns it into a simple ordinal IRT model with hierarchical predictors of ideal points.

The model was run in Stan with an iterative tuning procedure to get close to the posterior mode before tuning the Hamiltonian Monte Carlo parameters for the final sampling run. I ran 4 chains starting near the posterior mode for 50 tuning iterations followed by 100 sample draws, for a total of 400 points in the posterior.

The purpose of this model is to serve as a baseline for comparison with the salience model, so I will not deeply analyze the parameter estimates. I will highlight a few features that will be important for comparison, though.

The factor loadings in 5.4 show that most of the questions in the survey loaded primarily on the first latent dimension (the dimensions were rotated so that P1 had the most variation in loadings). There is still a good amount of spread along P2, though none of the questions loaded on P2 exclusively. The dimensions are scaled by individual-level variation in ideal points, so this scaling is partly because the ideal points along P2 had more individual-specific variation relative to the demographic effects.

This is somewhat visible in the graph of demographic predictors of ideal point in Figure 5.3, where many of the demographic effects lie close to 0 on the P2 dimension. There are exceptions, and particularly for Church, Income, BAgain, and Relig. These 4 variables along with Fem, MSing, and Educ will be discussed in more detail later. The lack of questions along P2 makes estimation of ideal points noisier along P2, as is clear in Figure 5.5.

5.4.2 Model with salience

The salience model was also run in Stan with an iterative tuning procedure to get close to the posterior mode before tuning the Hamiltonian Monte Carlo parameters for the final sampling run. I ran 4 chains starting near the posterior mode for 50

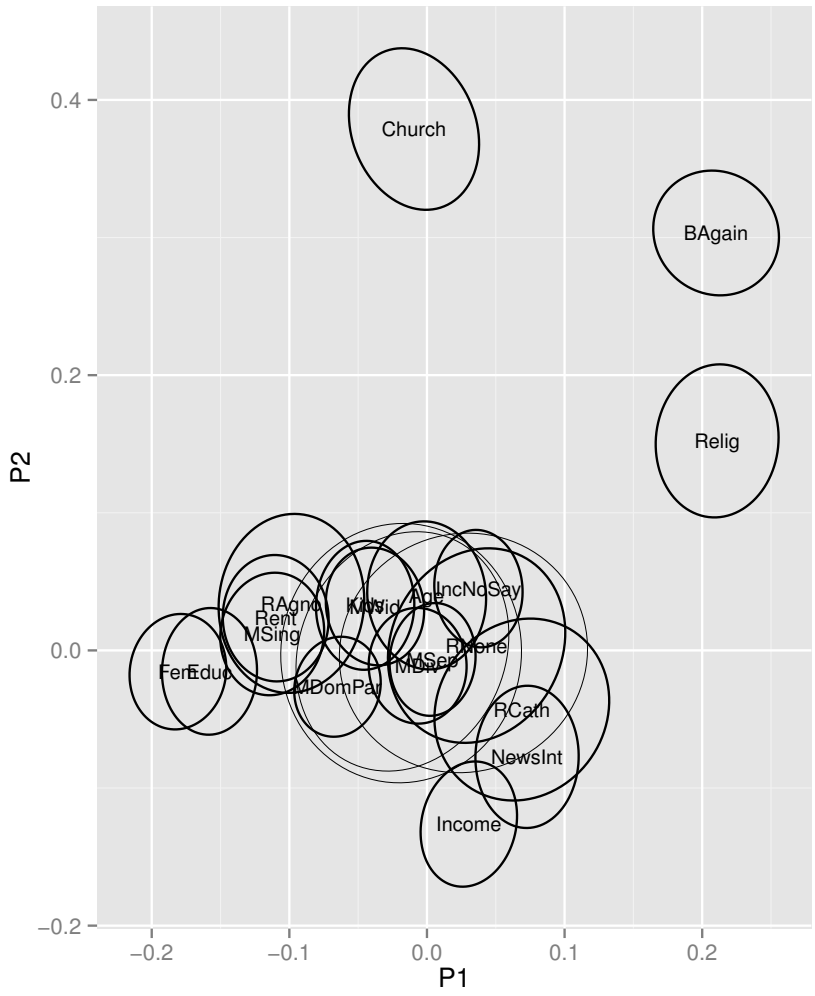


Figure 5.3: Effects of demographic parameters on political ideal points in model with uniform salience. Ellipses are the 40% confidence ellipse for the effect estimates (roughly comparable to 1 standard deviation).

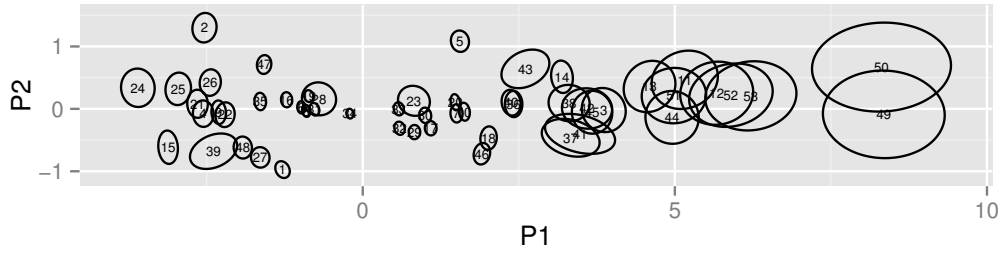


Figure 5.4: Loadings of the questions on each political dimension in the model with uniform salience. Ellipses are the 40% confidence ellipse for the effect estimates (roughly comparable to 1 standard deviation).

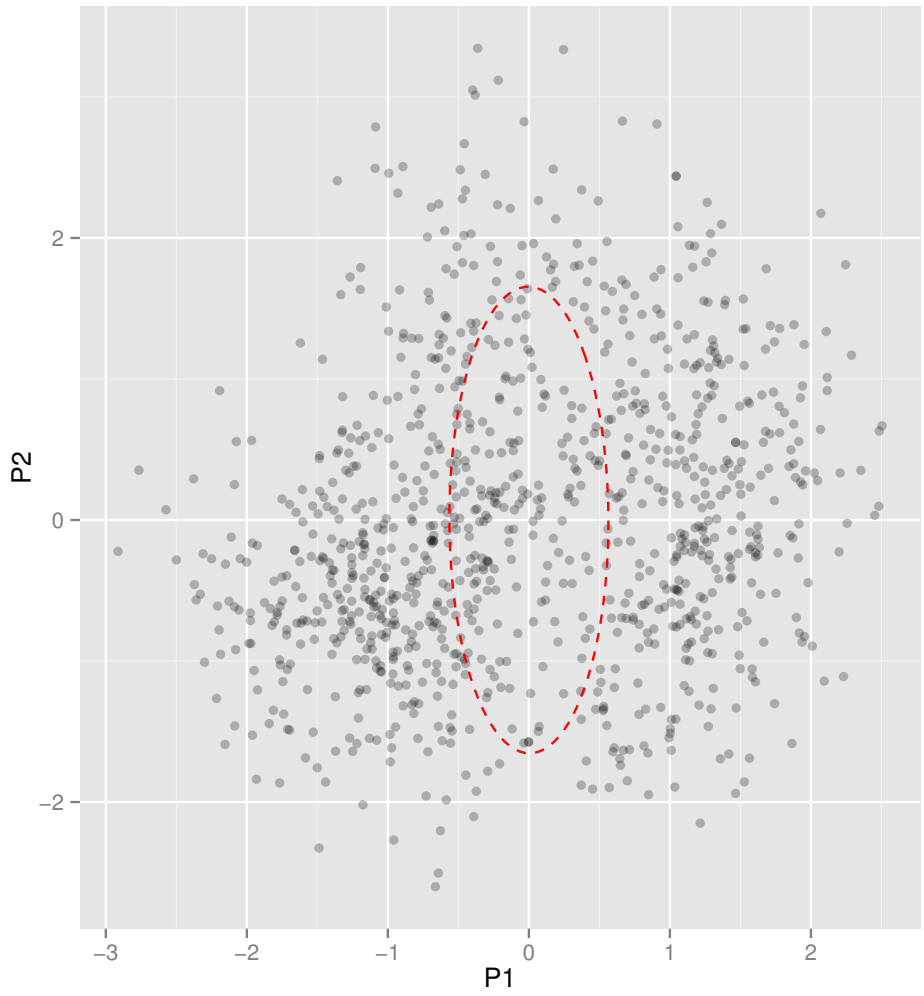


Figure 5.5: Distribution of a sample of ideal points (posterior mean for each individual) from the model with uniform salience. The red dotted ellipse is a 95% credible ellipse for a representative individual at the origin.

tuning iterations followed by 100 sample draws, for a total of 400 points in the posterior.

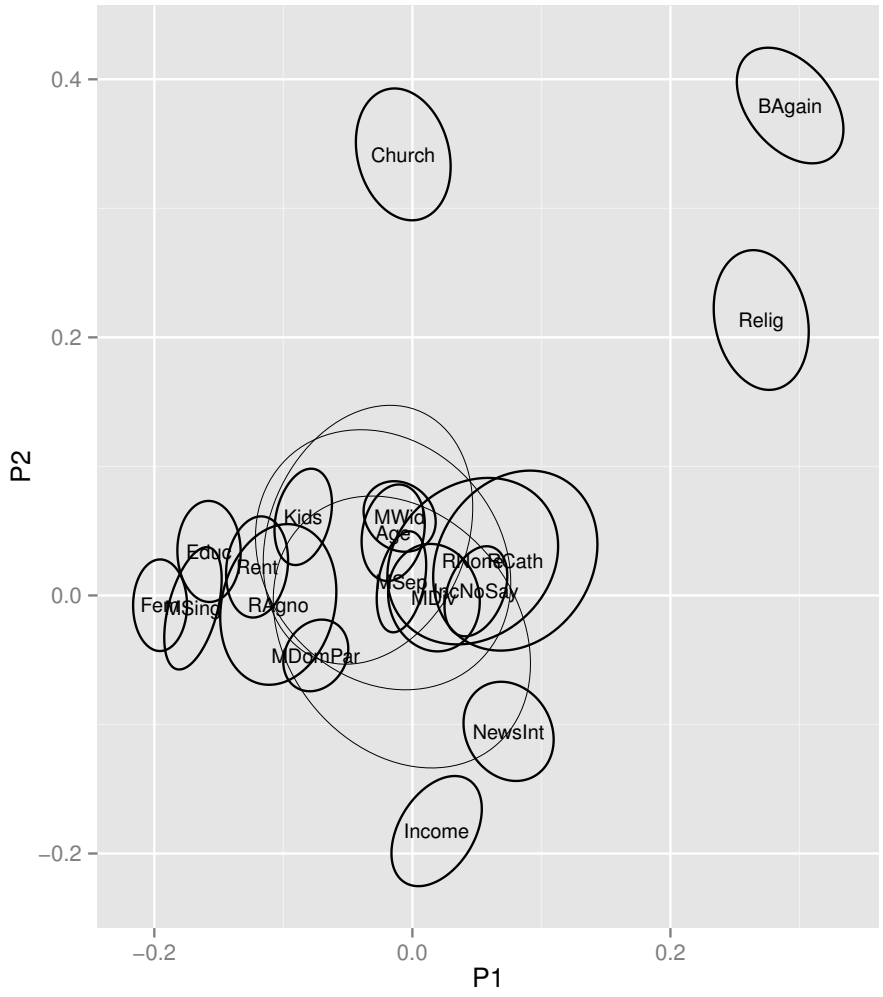


Figure 5.6: Effects of demographic parameters on political ideal points in the variable salience model. Ellipses are the 40% confidence ellipse for the effect estimates (roughly comparable to 1 standard deviation).

The salience model has somewhat different question loadings 5.7 than the uniform model. The handful of highly predictable questions high on the P1 scale are bent downwards somewhat to load on the P2 scale as well, though the ratio in this tail is still around 5:1 loading on P1:P2. Most of the other questions are in

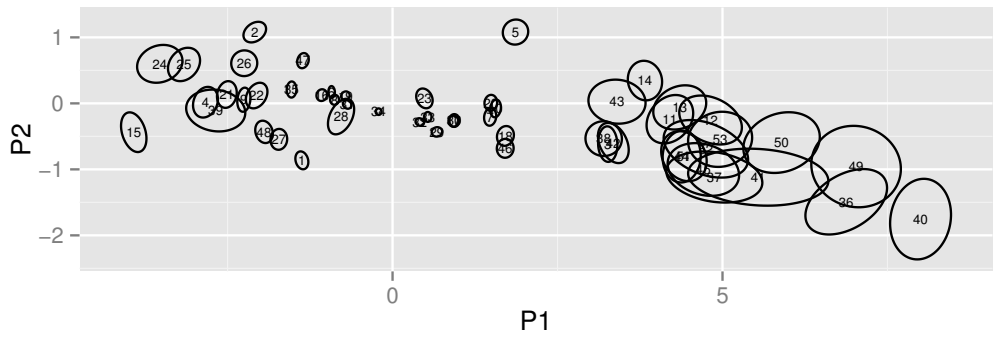


Figure 5.7: Loadings of the questions on each political dimension in the variable salience model. Ellipses are the 40% confidence ellipse for the effect estimates (roughly comparable to 1 standard deviation).

generally similar positions. There are a few loadings that shift around, particularly in questions like 24, 39, and 15 that had noisy salience estimates, but the overall picture is very similar aside from the bend at the high end of P1.

The parameter estimates (Figure 5.6) are also generally similar to those for the uniform model (Figure 5.3)—variables shift a bit, but none are switching directions or moving wildly. BAgain and Relig shift a bit to the right on P1, and MSing shifts slightly to the left, but there is very little other variation on P1. On P2, BAgain, Relig both become even more positive, while Church moves towards 0, and Income moves slightly more negative.

In terms of individual point estimates (Figure 5.8), the uncertainty on P1 is similar to before, while the uncertainty on P2 decreases slightly. The figure overstates the decrease slightly, though, since the variation in the mean P2 estimates increases as noise decreases, and so the scale of the P2 axis in Figure 5.8 is slightly expanded from that in Figure 5.5. The decrease is still there, it is just not quite so large as the figure makes it appear.

Now that we have seen that the changes to the latent space are relatively mild—demographic effects and question loadings are not wildly different from the previous model—we can move into the meat of the salience model: the demographic effects on salience, examples of individual salience ellipses and noise, and fitted indifference curves for individuals.

I visualize salience ellipses using level curves for a multivariate normal distribution with the salience matrix included as its covariance term. When an ellipse is stretched in one direction, the individual or group is impacted more by variation along that direction, and when the ellipse is compressed in a direction, they do not care as much about that variation. I find this more intuitive than the usual visualization of centered indifference curves, though indifference curves will be helpful when we are also visualizing individuals' ideal points.

The demographic effects are each small on their own, but they build on each other to produce a substantial amount of variation in individual salience estimates. For visualization, I look at the effects of having a values of 10 on each (centered

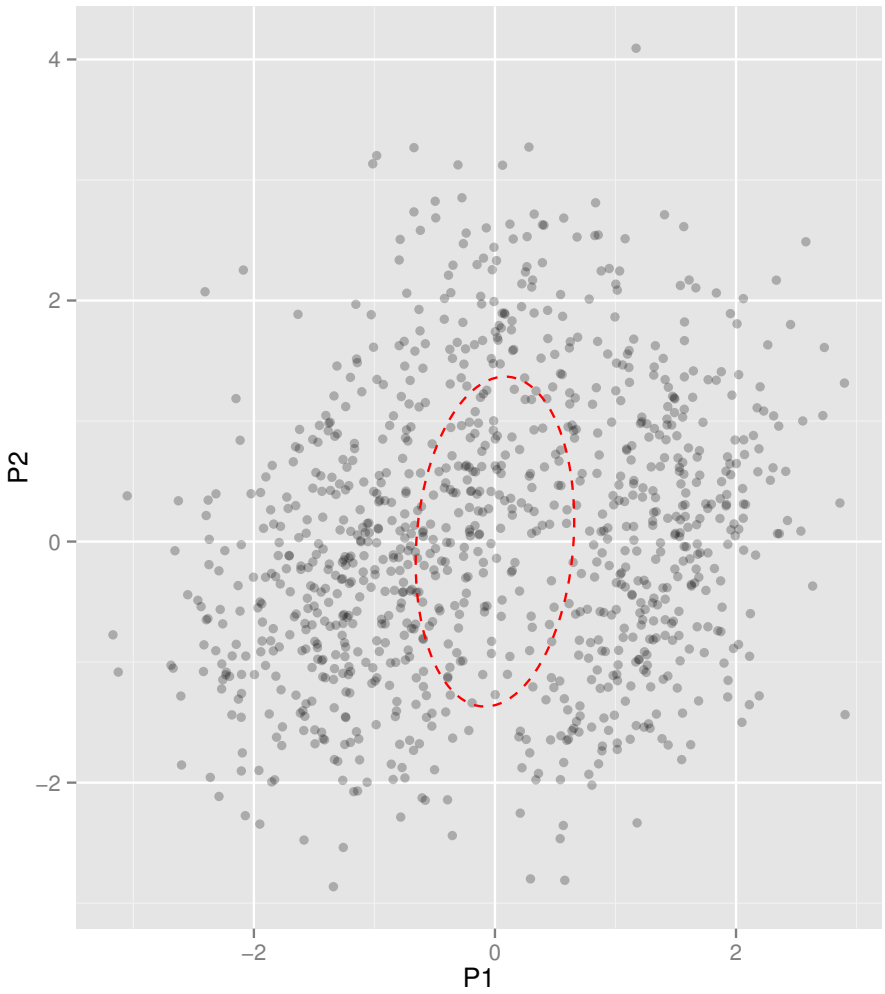


Figure 5.8: Distribution of a sample of ideal points (posterior mean for each individual) from the variable salience model. The red dotted ellipse is a 95% credible ellipse for a representative individual at the origin.

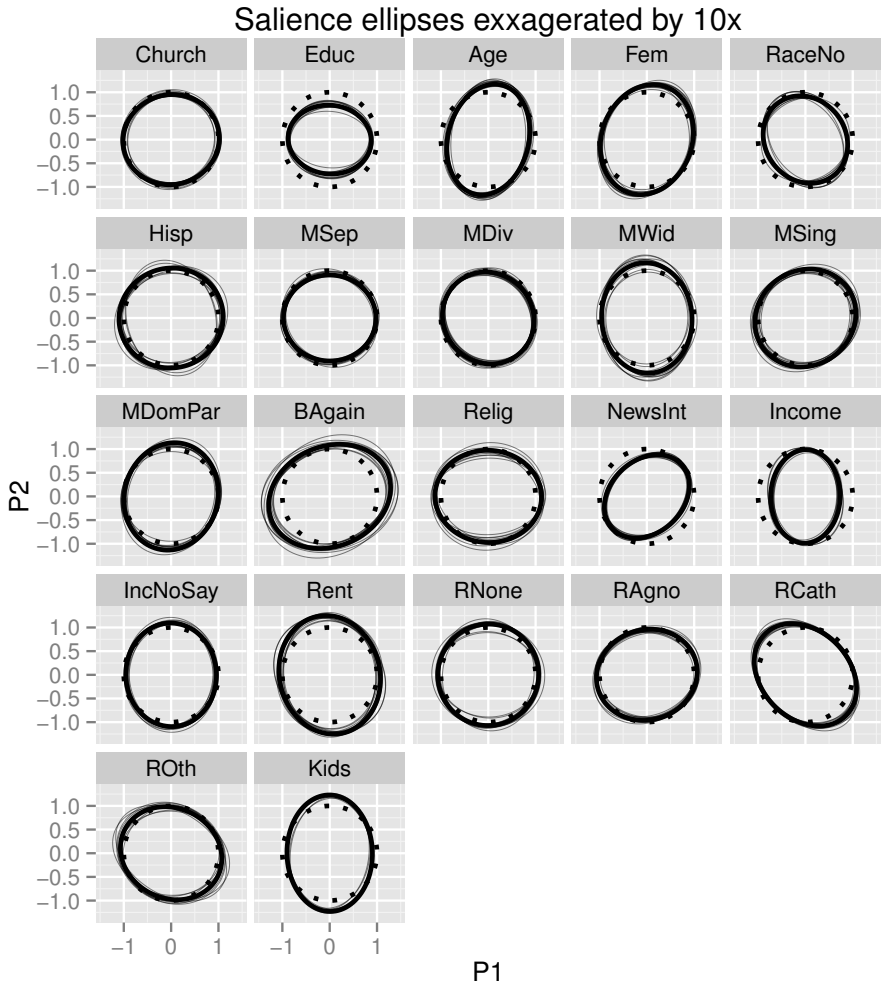


Figure 5.9: Salience ellipses for individuals with a value of 10 on each of the variables and 0 on all other demographics. The demographic variables are scaled to mean 0 and variance 1, so this extreme value is unrealistic and just helps with visualization. The dotted unit circle provides a 'no effect' reference.

and scaled) demographic variable and 0 on all other variables and compare that to a baseline of having 0 on all variables. Exaggerating the effect by 10 is necessary so that we can actually see variation in the ellipses. The plots show the mean ellipse and a sample of drawn ellipses to get a sense for the estimation uncertainty in the model. The mean ellipse is calculated by taking the pointwise average of all the log salience matrices, and then exponentiating the resulting matrix and constructing the ellipse.

There are a number of variables that impact the salience of P2 substantially. Age, Fem, MWid, Rent, Income, and Kids all increase the relative importance of P2, while Educ and Relig decrease it. The salience effects are not limited to the primary axes, so RCath and RaceNo produce greater salience in the diagonal (P2-P1) direction, while BAgain and NewsInt produce more relative salience on (P1+P2). Some factors like education and income decrease overall salience, while others like BAgain and Rent increase it.

The salience estimates are closely related to the demographic effects on ideal points. BAgain increases salience in a roughly (P1+P2) direction, and the BAgain ideal point effect is shifted out along this axis in the salience model compared to the uniform model. Salience impacts how parameters move between the uniform and salience models, but it is not strictly connected to the actual parameters.

One of the more interesting effects of salience occurs when the ideal point parameter and axes of the salience ellipse are roughly 45°. In party strategy models, this allows a candidate to capture more of a certain group by moving perpendicular to the group's ideal point. We see roughly this pattern for Relig. Relig shifts ideal points in a (P1+P2) direction, but it increases salience only along P1. So more religious people are in the upper right of the latent space, but care about movement along P1. A candidate at the origin could then boost their support among religious people by shifting slightly *downwards* to the right, which is closer along P1, but not closer in the latent space overall.

Similar effects can occur on cross terms, for example, between RCAth and Fem, where the salience impact of RCath and the ideal point impact of Fem are

neither parallel nor perpendicular. A campaign trying to target religious voters might want to tune their messages differently for males and females, since religious males and females can be swayed by different shifts. This is a huge simplification, as there are many other variables to consider, but it shows the utility of modeling salience at a group level.

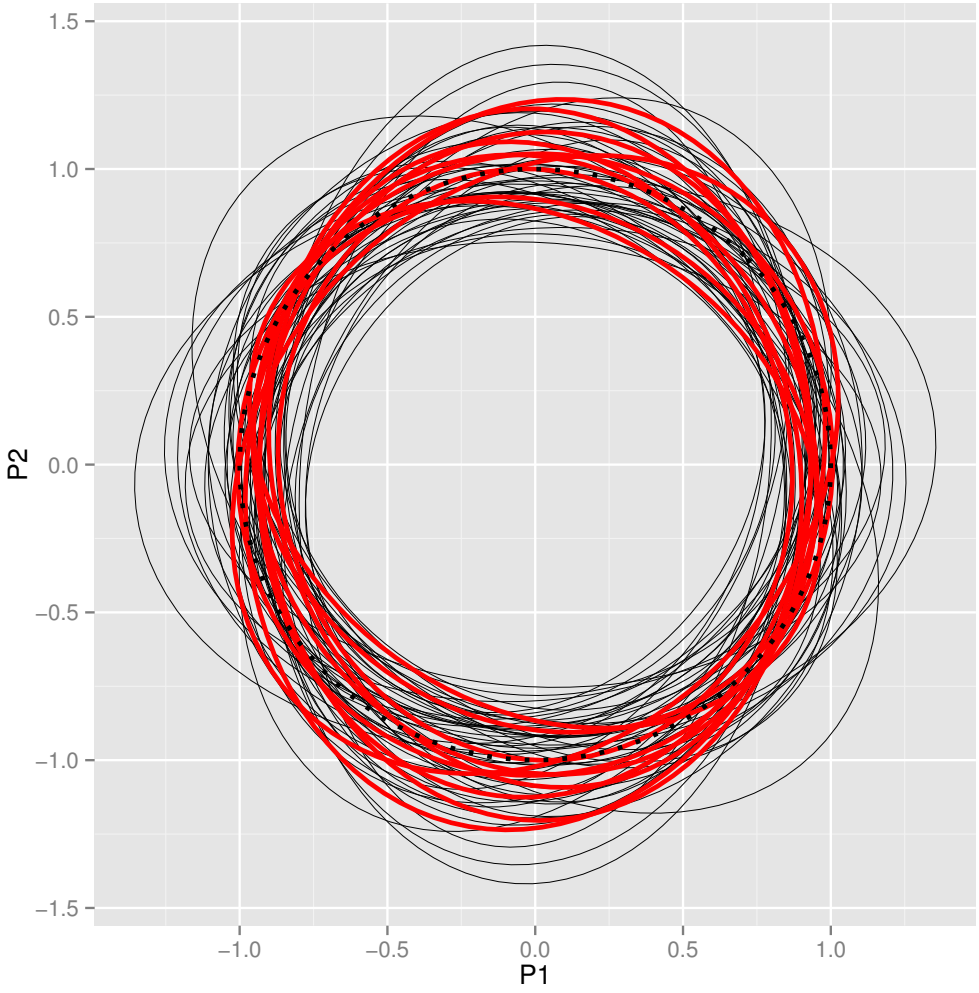


Figure 5.10: Example individual-specific salience noise for a sample of individuals (gray), and multiple draws from a single individual (red). The dotted unit circle provides a 'no effect' reference.

In terms of individual estimates, the model produces noisy but moderately

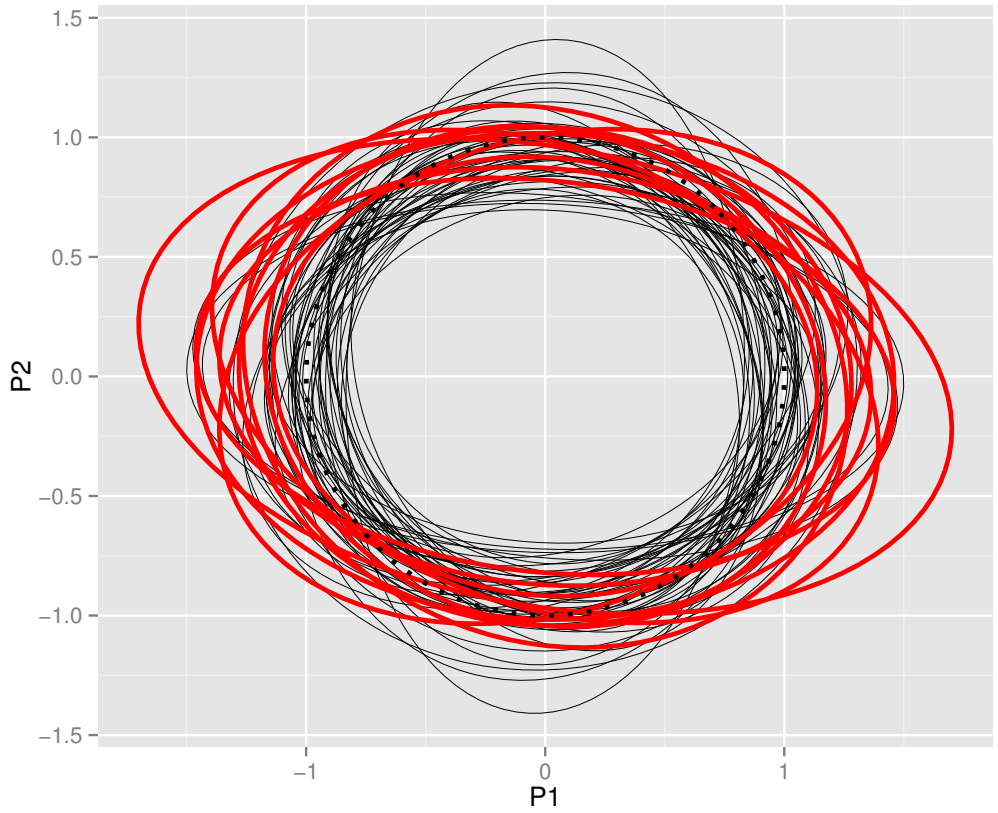


Figure 5.11: Example individual salience ellipses (combination of demographic factors and individual variation) for a sample of individuals (gray), and multiple draws from a single individual (red). The dotted unit circle provides a 'no effect' reference.

informative posteriors on the individual-specific salience noise. The red ellipses in Figure 5.11 for ellipses drawn for a single individual are tighter together than the thin black ellipses for a sample across individuals. Some of this is because the red ellipses include demographic information about the individual while the black ones are drawn from across the population, but when we strip away demographic effects and look just at an individual's noise term (Figure 5.10, for a different individual), the red ellipses are still tighter together than the black ones, and so we can estimate individual-specific variation in this model, albeit noisily.

Finally, to visualize how salience actually impacts the ideal points and indifference curves of voters, I plot a sample of mean ideal points and mean indifference curves—level curves for a multivariate normal distribution centered at the ideal point with covariance equal to the inverse of the salience matrix. The key result is that we see a substantial amount of variation in the salience across the population. Some of the ellipses are much larger than others, some are nearly circular while others are highly stretched, and throughout the distribution of ideal points, we see variation in salience. Individuals who have similar ideal points may still behave quite differently as a result of their salience differences.

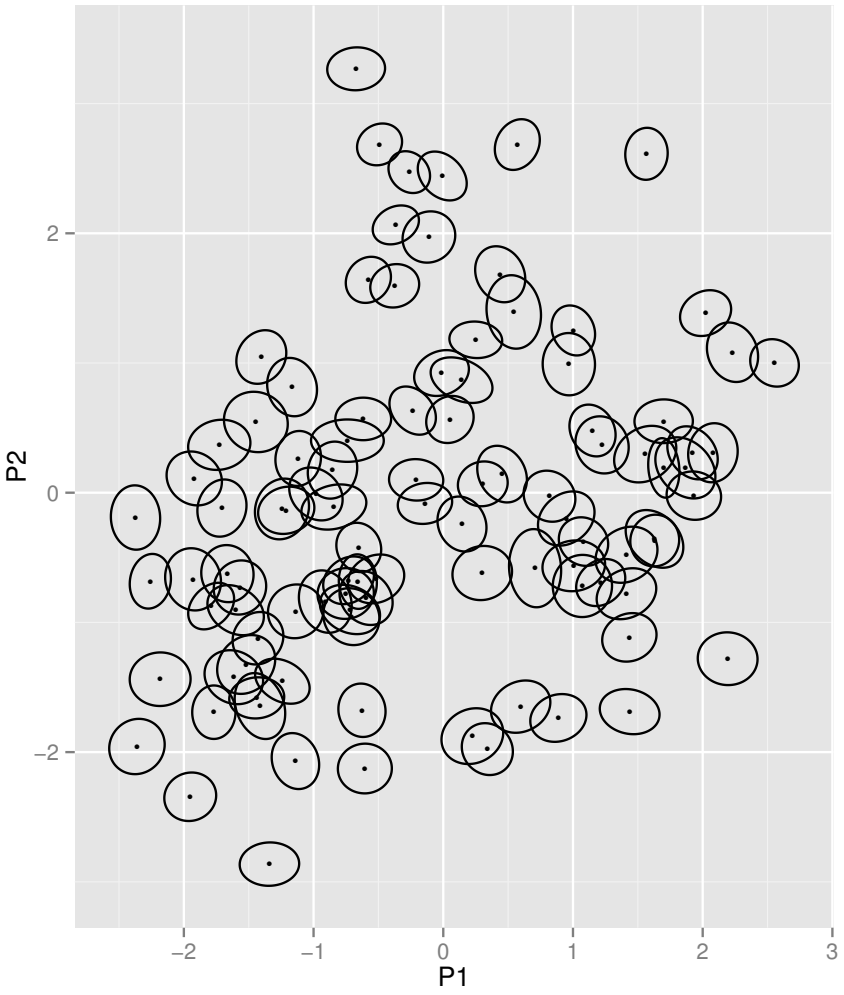


Figure 5.12: Mean ideal points and indifference curves for a sample of individuals from the variable salience model

5.5 Conclusions

The technique developed here offers a way to simultaneously estimate ideal points and salience on multiple dimensions from the types of survey data that we commonly have in political science, and in doing so, it opens up a new set of modeling options for researchers, particularly those doing experimental work related to political ideology.

Estimates of ideal points and salience can be input directly into formal models to understand where parties should be moving in order to maximize votes, or the demographic parameters can be used to simulate possible distributions of ideal points and salience independent of the pecularities of the survey sample used for estimation. This method provides the necessary inputs for actually testing formal models of party strategy under heterogeneous salience, and if the models survive the test, for making new predictions of party and voter behavior.

Thanks to the hierarchical design, researchers can include treatment variables in the model and estimate their impact on broad ideological dimensions rather than examining single issues in isolation or constructing ad-hoc scales that vary from paper to paper and lab to lab. They can examine how treatments change salience over broad dimensions like social or economic policies as a whole, rather than just looking at how direct manipulations on a specific dimension change specific outcomes.

In the world of political campaigns, changing how a voter thinks about an issue is sometimes easier than changing what they think about it—adjusting salience is easier than shifting an ideal point—and this method provides a way for campaigns and parties to test how their messages are impacting not only preferences, but how people translate those preferences into specific political choices.

Similarly, it is very likely that some biological and social factors operate by not directly altering preferences, but by changing what factors we consider: how likely we are to remember empathizing with a homeless person when voting on a tax increase to fund shelters, or instead to remember how the city's money

was wasted the last time we increased taxes. Those are impacts on salience, not direct impacts on preferences, and thus hierarchical estimation of salience makes it possible to understand what individual characteristics, genes, physiological factors, or experimental treatments lead people to care about certain issue dimensions. In looser terms, it helps us figure out why people think about politics differently from one another.

Appendices

Appendix A

CCES political questions

Questions that begin with 'CC3...' were asked before the 2010 election, and those that start with 'CC4...' were asked after. Questions CC322 and CC414 were split into binary variables that denoted whether respondents selected each option. Questions CC328 and CC329 were recoded into three variables to denote whether defense, domestic, or tax solutions were the best, middle, or worst options.

CC320 In general, do you feel that the laws covering the sale of firearms should be made more strict, less strict, or kept as they are?

- More Strict
- Kept As They Are
- Less Strict

CC321 From what you know about global climate change or global warming, which one of the following statements comes closest to your opinion?

- Global climate change has been established as a serious problem, and immediate action is necessary.
- There is enough evidence that climate change is taking place and some action should be taken.
- We don't know enough about global climate change, and more research is necessary before we take any actions.

- Concern about global climate change is exaggerated. No action is necessary.
- Global climate change is not occurring; this is not a real issue.

CC322 What do you think the U.S. government should do about immigration? Select all that apply.

- Grant legal status to all illegal immigrants who have held jobs and paid taxes for at least 3 years, and not been convicted of any felony crimes.
- Increase the number of border patrols on the US-Mexican border.
- Allow police to question anyone they think may be in the country illegally.

CC324 Which one of the opinions on this page best agrees with your view on abortion?

- By law, abortion should never be permitted
- The law should permit abortion only in case of rape, incest or when the woman's life is in danger
- The law should permit abortion for reasons other than rape, incest, or danger to the woman's life, but only after the need for the abortion has been clearly established
- By law, a woman should always be able to obtain an abortion as a matter of personal choice

CC325 Some people think it is important to protect the environment even if it costs some jobs or otherwise reduces our standard of living. Other people think that protecting the environment is not as important as maintaining jobs and our standard of living. Which is closer to the way you feel, or haven't you thought much about this?

- Much more important to protect environment even if lose jobs and lower standard of living

- Environment somewhat more important
- About the same
- Economy somewhat more important
- Much more important to protect jobs, even if environment worse

CC326 Do you support a Constitutional Amendment banning Gay Marriage?

- Yes
- No

CC327 Affirmative action programs give preference to racial minorities in employment and college admissions in order to correct for past discrimination. Do you support or oppose affirmative action?

- Strongly support
- Somewhat support
- Somewhat oppose
- Strongly oppose

CC328 The federal budget deficit is approximately $600 billion this year. If the Congress were to balance the budget it would have to consider cutting defense spending, cutting domestic spending (such as Medicare and Social Security), or raising taxes to cover the deficit. What would you most prefer that Congress do - cut domestic spending, cut defense spending, or raise taxes?

- Cut Defense Spending
- Cut Domestic Spending
- Raise Taxes

CC329 What do you least want Congress to do?

- Cut Defense Spending
- Cut Domestic Spending
- Raise Taxes

CC332A Congress considered many important bills over the past two years. For each of the following tell us whether you support or oppose the legislation in principle. American Recovery and Reinvestment Act–Authorizes $787 billion in federal spending to stimulate economic growth in the US.

- Support
- Oppose

CC332B State Children's Health Insurance Program–Program insures children in low income households. Act would renew the program through 2014 and include 4 million additional children.

- Support
- Oppose

CC332C American Clean Energy and Security Act–Imposes a cap on carbon emissions and allow companies to trade allowances for carbon emissions. Funds research on renewable energy.

- Support
- Oppose

CC332D Comprehensive Health Reform Act–Requires all Americans to obtain health insurance. Allows people to keep current provider. Sets up health insurance option for those without coverage. Increases taxes on those making more than $280,000 a year.

- Support
- Oppose

CC332E Appoint Elena Kagan to the US Supreme Court

- Support
- Oppose

CC332F Financial Reform Bill–Protects consumers against abusive lending. Regulates high risk investments known as derivatives. Allow government to shut down failing financial institutions.

- Support
- Oppose

CC332G End Don't Ask, Don't Tell–Would allow gays to serve openly in the armed services.

- Support
- Oppose

CC332I Embryonic Stem Cell Research–Allow federal funding of embryonic stem cell research.

- Support
- Oppose

CC414 Would you approve of the use of U.S. military troops in order to. . . ? (Please check all that apply)

- Ensure the supply of oil
- Destroy a terrorist camp
- Intervene in a region where there is genocide or a civil war
- Assist the spread of democracy ? Protect American allies under attack by foreign nations
- Help the United Nations uphold international law

CC422a The Irish, Italians, Jews and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favors.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

CC422b Generations of slavery and discrimination have created conditions that make it difficult for Blacks to work their way out of the lower class.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

Appendix B

Proofs related to salience models

B.1 Equivalence of traditional and hyperplane-based constructions

The two model constructions for ordinal IRT with multiple latent dimensions are

- Traditional

$$y^* = \phi\lambda + \epsilon, \qquad \epsilon \sim \text{Normal}(0, 1) \tag{B.1}$$

$$y = 1 + \sum_j \{y^* < \eta_j\} \tag{B.2}$$

- Hyperplane-based

$$y^* = \left(\phi + \frac{\epsilon}{\|\lambda\|}\right)\lambda, \qquad \epsilon \sim \text{MVN}(0, I) \tag{B.3}$$

$$y = 1 + \sum_j \{y^* < \eta_j\} \tag{B.4}$$

In B.3, note that variation in ϵ perpendicular to λ has no impact on y^*. Since the distribution of ϵ is spherical, we can rotate it into a space that has $\lambda/\|\lambda\|$ as the first basis vector, call the rotated variable ϵ', and just look at the marginal distribution along that $\lambda/\|\lambda\|$ dimension. The transformed mean and variance of ϵ' are still 0 and I, and since the marginal distribution of any dimension of a

standard multivariate normal is simply a standard normal, so the distribution of $\epsilon' \cdot \frac{\lambda}{\|\lambda\|}$ is Normal$(0, 1)$, and so if we define a final variable $\epsilon'' = \epsilon' \cdot \frac{\lambda}{\|\lambda\|}$, that is equivalent to the ϵ term in B.1.

B.2 Choice model invariance to perpendicular shifts under uniform salience

We begin with a spatial choice model parameterized in terms of outcome policy positions with ideal point ϕ, question direction vector λ, question- and individual-specific noise ϵ, individual-specific salience matrix A, and cutpoints ζ.

$$y^* = \left(\phi + \alpha + \frac{\epsilon}{\|\lambda\|}\right)\lambda, \qquad \epsilon \sim \text{MVN}(0, I) \tag{B.5}$$

$$y = 1 + \sum_j \{y^* < \eta_j\} \tag{B.6}$$

and without loss of generality, assume that $\alpha \cdot \lambda = 0$ so that α captures shifts perpendicular to the question direction vector.

Since $\alpha\lambda = 0$, the manipulations to B.5 that make α disappear are trivial.

$$y^* = \phi\lambda + \alpha\lambda + \frac{\epsilon\lambda}{\|\lambda\|} \tag{B.7}$$

$$= \phi\lambda + \frac{\epsilon\lambda}{\|\lambda\|} \tag{B.8}$$

$$= \left(\phi + \frac{\epsilon}{\|\lambda\|}\right)\lambda \tag{B.9}$$

And so the distribution of observable outcomes y conditional on ϕ and λ are not affected by shifts in ϕ perpendicular to λ.

Bibliography

J.R. Alford, C.L. Funk, and J.R. Hibbing. Are Political Orientations Genetically Transmitted? *The American Political Science Review*, 99(2):153–167, May 2005.

J.R. Alford, J.R. Hibbing, N.G. Martin, and L.J. Eaves. Is There a" Party" in Your Genes? *Political Research Quarterly*, September 2009.

R.M. Alvarez and J. Brehm. Speaking in two voices: American equivocation about the Internal Revenue Service. *American Journal of Political Science*, 42(2):418–452, 1998.

R.M. Alvarez and J. Brehm. *Hard Choices, Easy Answers : Values, Information, and American Public Opinion*. Princeton University Press, 2002.

R.M. Alvarez and J. Nagler. Economics, Issues and the Perot Candidacy: Voter Choice in the 1992 Presidential Election. *American Journal of Political Science*, 39(2):714–744, August 1995.

R.M. Alvarez and J. Nagler. Economics, Entitlements, and Social Issues: Voter Choice in the 1996 Presidential Election. *American Journal of Political Science*, 42(4):1349–1363, October 1998.

R.M. Alvarez, J. Nagler, and S. Bowler. Issues, Economics, and the Dynamics of Multiparty Elections: The British 1987 General Election. *The American Political Science Review*, 94(1):131–149, March 2000.

J. Bafumi. A New Partisan Voter. *The Journal of Politics*, 71(01):1–24, January 2009.

M. Benedek and C. Kaernbach. A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, 190:80–91, 2010.

M. Benedek and C. Kaernbach. Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology*, 47:647–658, 2010.

H.E. Brady. Attitude Attribution: A Group Basis for Political Reasoning. *The American Political Science Review*, 79(4):1061–1078, December 1985.

D.H. Brainard. The Psychophysics Toolbox. *Spatial Vision*, 10(4):433–436, 1997.

K.A. Brownley, B.E. Hurwitz, and N. Schneiderman. Cardiovascular Psychophysiology. In *Handbook of Psychophysiology*, pages 224–264. Cambridge University Press, 2000.

E.A. Butler, F.H. Wilhelm, and J.J. Gross. Respiratory sinus arrhythmia, emotion, and emotion regulation during social interaction. *Psychophysiology*, 43(6):612–622, November 2006.

A. Campbell, P. Converse, W. Miller, and D. Stokes. *The American Voter*. The University of Chicago Press, Chicago and London, 1960.

J. Clinton, S. Jackman, and D. Rivers. The Statistical Analysis of Roll Call Data. *The American Political Science Review*, 98(2):355–370, 2004.

P.E. Converse. *The Nature of Belief Systems in Mass Publics*. In D.E. Apter, editor, *Ideology and Discontent*. Free Press, New York, 1964.

H.A. Demaree, J.L. Robinson, D.E. Everhart, and B.J. Schmeichel. Resting RSA is associated with natural and self-regulated responses to negative emotional stimuli. *Brain and Cognition*, 56(1):14–23, October 2004.

S.I. Donaldson and E.J. Grant-Vallone. Journal of Business and Psychology, Volume 17, Number 2 - SpringerLink. *Journal of Business and Psychology*, 17(2):245–260, 2002.

S. Feldman. Structure and consistency in public opinion: The role of core beliefs and values. *American Journal of Political Science*, 32(2):416–440, 1988.

R. Ferber. Item Nonresponse in a Consumer Survey. *Public Opinion Quarterly*, 30(3):399, 1966.

M.P. Fiorina. Economic Retrospective Voting in American National Elections: A Micro-Analysis. *American Journal of Political Science*, 22(2):426–443, May 1978.

J.H. Fowler and C.T. Dawes. Two Genes Predict Voter Turnout. *The Journal of Politics*, 70(3):579–594, 2008.

J. Francis and L. Busch. What We Now Know about" I Don't Knows". *The Public Opinion Quarterly*, 39(2):207–218, 1975.

A.S. Gerber, G.A. Huber, and E. Washington. Party Affiliation, Partisanship, and Political Beliefs: A Field Experiment. *The American Political Science Review*, 104(4):720–744, 2010.

G. Gigerenzer and D.G. Goldstein. Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103(4):650–669, 1996.

G. Glasgow. Mixed Logit Models for Multiparty Elections. *Political Analysis*, 9(1):116–136, 2001.

J. Graham, B.A. Nosek, J. Haidt, and R. Iyer. Mapping the moral domain. *Journal of Personality and Social Psychology*, 101(2):366–385, 2011.

P. Grossman, J. Beek, and C. Wientjes. A Comparison of Three Quantification Methods for Estimation of Respiratory Sinus Arrhythmia. *Psychophysiology*, 27(6):702–714, November 1990.

J.J. Heckman and J.M. Snyder, Jr. Linear probability models of the demand for attributes with an empirical application to .... *Rand Journal of Economics*, 28(0):S142–S189, January 1997.

M.J. Hinich and M.C. Munger. Analytical Politics. Cambridge University Press, April 1997.

M.D. Hoffman and A. Gelman. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *arXiv*, November 2011.

L. Huddy, S. Feldman, C. Taber, and G. Lahav. Threat, anxiety, and support of antiterrorism policies. *American Journal of Political Science*, 49(3):593–608, July 2005.

S. Iyengar, K. Hahn, and J. Krosnick. Selective exposure to campaign communication: The role of anticipated agreement .... *The Journal of Politics*, January 2008.

S. Jackman. Multidimensional analysis of roll call data via Bayesian simulation: Identification, estimation, inference, and model checking. *Political Analysis*, 2001.

M.K. Jennings and R.G. Niemi. The Persistence of Political Orientations: An Over-Time Analysis of Two Generations. *British Journal of Political Science*, 8(03):333–363, January 2009.

R. Kanai, T. Feilden, and C. Firth. Political orientations are correlated with brain structure in young adults. *Current Biology*, (21):677–680, April 2011.

K. Keissar, L.R. Davrath, and S. Akselrod. Coherence analysis between respiration and heart rate variability using continuous wavelet transform. *Phil. Trans. R. Soc. A*, 367:1393–1406, 2009.

P.J. Lang, M.M. Bradley, and B.N. Cuthbert. International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. Technical report, University of Florida, Gainesville, FL, 2008.

B.E. Lauderdale. Unpredictable Voters in Ideal Point Estimation. *Political Analysis*, 2010.

J. Leigh and C.R. Martin, Jr. " Don't Know" Item Nonresponse in a Telephone Survey: Effects of Question Form and Respondent Characteristics. *Journal of Marketing Research*, pages 418–424, 1987.

G. Loosveldt, J. Pickery, and J. Billiet. Item nonresponse as a predictor of unit nonresponse in a panel survey. *Journal of Official Statistics*, 18(4):545–558, 2002.

A. Martin and K. Quinn. Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999. *Political Analysis*, 10(2):134–153, 2002.

A.D. Martin, K.M. Quinn, and J.H. Park. MCMCpack: Markov Chain Monte Carlo in R. *Journal of Statistical Software*, 2011.

K. Mattes, M. Spezio, H. Kim, A. Todorov, R. Adolphs, and R.M. Alvarez. Predicting Election Outcomes from Positive and Negative Trait Assessments of Candidate Images. *Political Psychology*, 31(1):41–58, 2010.

J.A. Mikels, B.L. Fredrickson, G.R. Larkin, C.M. Lindberg, S.J. Maglio, and P.A. Reuter-Lorenz. Emotional category data on images from the International Affective Picture System. *Behavior research methods*, 37(4):626–630, November 2005.

C. Oveis, A.B. Cohen, J. Gruber, M.N. Shiota, J. Haidt, and D. Keltner. Resting respiratory sinus arrhythmia is associated with tonic positive emotionality. *Emotion*, 9(2):265–270, 2009.

D. Oxley, K. Smith, J.R. Alford, M. Hibbing, J. Miller, P. Hatemi, and J.R. Hibbing. Political Attitudes Vary with Physiological Traits. *Science (New York, NY)*, 321(5896):1667, September 2008.

M. Peress. Small Chamber Ideal Point Estimation. *Political Analysis*, 17(3):276, July 2009.

K.T. Poole and H. Rosenthal. A Spatial Model for Legislative Roll Call Analysis. *American Journal of Political Science*, 29(2):357–384, May 1985.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012.

D. Rivers. Heterogeneity in models of electoral choice. *American Journal of Political Science*, January 1988.

D. Rivers. Identification of Multidimensional Spatial Voting Models. Technical report, July 2003.

A. Rosenblatt, J. Greenberg, S. Solomon, T. Pyszczynski, and et al. Evidence for terror management theory: I. The effects of mortality salience on reactions to those who violate or uphold cultural values. *Journal of Personality and Social Psychology*, 57(4):681–690, 1989.

J. Rottenberg, R.D. Ray, and J.J. Gross. Emotion Elicitation Using Films. *The handbook of emotion elicitation and assessment*, pages 9–28, 2007.

D. Rubin, H. Stern, and V. Vehovar. Handling "don't know" survey responses: the case of the Slovenian plebiscite. *Journal of the American Statistical Association*, pages 822–828, 1995.

S. Skaperdas and B. Grofman. Modeling Negative Campaigning. *The American Political Science Review*, 89(1):49–61, March 1995.

P.M. Sniderman and J. Bullock. *A Consistency Theory of Public Opinion and Political Choice: The Hypothesis of Menu Dependence in Studies*. In W.E. Saris and P.M. Sniderman, editors, *Public Opinion: Attitudes, Nonattitudes, Measurement Error, and Change*, pages 338–353. Princeton University Press, 2004.

M.L. Spezio, A. Rangel, R.M. Alvarez, J.P. O'Doherty, K. Mattes, A. Todorov, H. Kim, and R. Adolphs. A neural basis for the effect of candidate appearance

on election outcomes. *Social Cognitive and Affective Neuroscience*, pages 1–9, October 2008.

Stan Development Team. *Stan: A C++ Library for Probability and Sampling*, 1.3 edition, 2013.

Stan Development Team. *Stan Modeling Language*, 1.3.0 edition, 2013.

A. Todorov, S. Chaiken, and M.D. Henderson. The Heuristic-Systematic Model of Social Information Processes. In *The Persuasion Handbook*. SAGE, 2002.

J.E. Transue. Identity Salience, Identity Acceptance, and Racial Policy Attitudes: American National Identity as a Uniting Force. *American Journal of Political Science*, 51(1):78–91, January 2007.

S. Verba, K.L. Schlozman, and H.E. Brady. *Voice and equality*. civic voluntarism in American politics. Harvard Univ Pr, 1995.

J. Zaller and S. Feldman. A simple theory of the survey response: answering questions versus revealing preferences. *American Journal of Political Science*, January 1992.