Supplementary Information for "Moderates"

Anthony Fowler, Seth J. Hill, Jeffrey B. Lewis, Chris Tausanovitch, Lynn Vavreck, and Christopher Warshaw

A EM algorithm for Issue Opinion Model Estimation

In this appendix, we describe the EM algorithm (Dempster, Laird, and Rubin, 1977) used to estimate the parameters of the likelihood function shown in Equation 5. Following the notation introduced in in the main text and for convenience letting $\theta = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\lambda})$, we begin by forming the "complete data" log likelihood,

$$\ell(\theta) = \sum_i \sum_t v_{it} \log L_t(\boldsymbol{y}_{i\cdot}; \theta)$$

where $v_{it} = 1$ if the ith respondent is of type $t \in \{1, 2, 3\}$ and 0 otherwise. Note that $\sum_t v_{it} = 1$ for all respondents i. In the complete data problem, the type of each respondent is known and indicated by v. Of course, v is not observable. However, the EM algorithm is formed by iteratively maximizing the expected value of ℓ over the unknown values of v given estimates θ and the observed data.

In particular, we form the expectation of ℓ over v as

$$\begin{aligned} Q(\theta|\theta^{(s)}) &= \textstyle\sum_i \sum_t \mathrm{E}_{v_{it}|\boldsymbol{y_{i\cdot}},\theta^{(s)}} \left(v_{it} \log L_t(\boldsymbol{y}_{i\cdot}; \theta) \right) \\ &= \textstyle\sum_i \sum_t \mathrm{E}_{v_{it}|\boldsymbol{y_{i\cdot}},\theta^{(s)}} \left(v_{it} \right) \log L_t(\boldsymbol{y}_{i\cdot}; \theta) \\ &= \textstyle\sum_i \sum_t w_{it} \log L_t(\boldsymbol{y}_{i\cdot}; \theta) \end{aligned}$$

where

$$w_{it} = \frac{\bar{w}_t^{(s)} L_t(\boldsymbol{y}_{i\cdot}; \theta^{(s)})}{\sum_{t'} \bar{w}_{t'}^{(s)} L_{t'}(\boldsymbol{y}_{i\cdot}; \theta^{(s)})}$$

and $s = 0, 1, 2, \ldots$ indicates the current step of the EM algorithm.

A.1 The EM algorithm

The algorithm proceeds as follows:

1. The step counter, s, is set to zero and start values for $\theta^{(0)}$ and $\bar{w}_t^{(0)}$ for $t = 1, 2, 3$ are selected.

2. *E-Step*: w_{it} is formed for all i and t given $\theta^{(s)}$ and $\bar{w}_t^{(s)}$.

3. *M-Step*: Q is maximized in three parts yielding $\theta^{(s+1)}$ and $\bar{w}_t^{(s+1)}$ for $t = 1, 2, 3$. These three parts are as follows:

   a. For the parameters describing the issue opinions of respondents of type 1,

   $$\sum_i w_{i1} \log L_1(y_{i\cdot}; \boldsymbol{\alpha}, \boldsymbol{\beta})$$

   is maximized to update the estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

   b. The parameters describing the issue opinion of respondents of type 2 are updated as weighted means, $\lambda_j^{(s+1)} = \frac{\sum_{i \in \mathcal{N}_j} w_{i2} y_{ij}}{\sum_{i \in \mathcal{N}_j} w_{i2}}$ for $j = 1, \ldots, J$ where $\mathcal{N}_j$ is the set of respondents who answered question j.

   c. The sample proportions belonging to each type are updated as $\bar{w}_t^{(s+1)} = \sum_i w_{it}/N$ for $t = 1, 2, 3$.

4. s is incremented and the process repeated from (2) until convergence.

E-step details: As shown above, the calculation of w_{it}, requires the evaluation of L_1, L_2 and L_3. The likelihood of individual i's issue question responses if he is of type 3, $L_3(\boldsymbol{y}_{i\cdot})$, is simply a function of the number of responses given, does not depend on $\theta^{(s)}$, and is straightforward to calculate using Equation 2. Similarly, the calculation of the likelihood of individual i's issue question responses if she is of type 2, $L_2(\boldsymbol{y}_{i\cdot}, \boldsymbol{\lambda})$, requires only the straight-forward application of Equation 3.

The calculation of the likelihood of individual i's issue question responses if she is of type 1, $L_1(\boldsymbol{y}_{i\cdot}, \boldsymbol{\alpha}, \boldsymbol{\beta})$, is more complicated because it involves the calculation of the integral shown in Equation 1 as well as an estimate of the distribution of ideal points, f. We approximate f and the integral using Monte Carlo methods. In particular, we draw a sample from the current estimated ideal points, $\dot{x}_k$ for $k = 1, \ldots, M$, of size M. The sample is drawn independently and with replacement with sampling weights that are proportional to the current weights, w_{i1} for $i = 1, \ldots, N$. Because the estimated ideal points are drawn in proportion of the type 1 membership weights, the resulting sample is (approximately) drawn from f. Given this Monte Carlo draw from f, the integral in Equation 1 is approximated as

$$L_1(\boldsymbol{y}_{i\cdot}; \boldsymbol{\alpha}, \boldsymbol{\beta}) \approx \sum_{k=1}^{M} \prod_{j \in \mathcal{J}_i} \Lambda\left(\beta_j(\dot{x}_k - \alpha_j)\right)^{y_{ij}} \left(1 - \Lambda\left(\beta_j(\dot{x}_k - \alpha_j)\right)\right)^{1-y_{ij}}.$$

M-Step details: As part of the M-step, $\sum_i w_{i1} \log L_1(y_{i\cdot}; \boldsymbol{\alpha}, \boldsymbol{\beta})$ is maximized to update the estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. These estimates are arrived at using a weighted version of the quadratic majorization approach of de Leeuw (2011b) where the weights are w_{i1} for $i = 1, \ldots, N$. Each ideal point is estimated as a fixed effect. Thus, the distribution of ideal points is estimated non-parametrically in this approach. This is also equivalent to a weighted version of the spatial voting model estimation method described in Imai, Lo, and Olmsted (2016).

B Power Simulations and Data Selection

The datasets that we use in this paper have very large numbers of respondents. However they have fewer policy questions than we would like, particularly for a model of this level of complexity. So it is important to assess the power of the model with respect to the number of items.

In Section B.1 of this appendix, we show two sets of simulations that examine how many items in a survey are necessary to accurately estimate both respondents' type and their spatial ideal points. Overall, this analysis leads us to conclude that about 20 policy items are necessary to accurately estimate all of the parameters in the model that we present in the main text.

If an analyst used less than 20 items:

- The respondents' one-dimensional ideal points would be estimated somewhat less accurately (see the left panels of Figures A1 and A2).
- More problematically, the respondents' types (Downsian, Conversion, Inattentive) would be estimated substantially less accurately when there are fewer than 20 items (see the middle panels of Figures A1 and A2), and estimates of the overall composition of the sample between these types would be greatly biased (right panels of Figures A1 and A2). Indeed, the accuracy of the estimated types and the overall composition of the sample between types increases dramatically at around 20 items.

In section B.2 of this appendix, we examine other features of the policy items that can increase the accuracy of the model parameters. Here too, we consistently find that the number of items is the most important predictor of model accuracy.

B.1 Simulations on effect of the number of policy items on model accuracy

There is no simple power calculation that will tell us how many items we need to get precise estimates of our model parameters. In the absence of such a formula we use simulations. We simulate our model two ways, both using actual data. First, we run the model on an existing data set, randomly selecting among the available items for each trial. We vary the number of items from 10 to the full number, in this case 32, conducting three trials for each number of items. In each trial we estimate the parameters of the model. We compare these estimates to the estimates we obtain using all 32 items.

Our second simulation method takes the estimated parameters from the full dataset and uses them to simulate new datasets. On each trial we randomly select a number of items M, doing this three times for each of M in 10 to 32, as before. Then we simulate a dataset using the estimated parameters from the full model for those items, and estimate our model on this simulated dataset. We continue to use the parameters estimated using all 32 items as our benchmark.

The first method has the advantage that it does not assume that our model is correctly specified. It simply takes a real dataset and estimates the parameters for various numbers of items. However this method is susceptible to the possibility that our conclusions will be affected by the idiosyncrasies of the dataset we choose. The second method assumes that our model *is* correctly specified. The data simply provide a set of plausible parameters to use for the simulations. The conclusions using this method are more generalizable in the sense that they should capture cases where there is similar heterogeneity in the parameters and the model is appropriate.

Among the datasets available to us, the 2014 CCES had the greatest number of items at the time of this simulation analysis.[1] The parameters of interest to us are the estimated type probabilities and ideal points for Downsian types. We want to know when we can make

[1] We later added the 2015, 2017 and 2018 CCES.

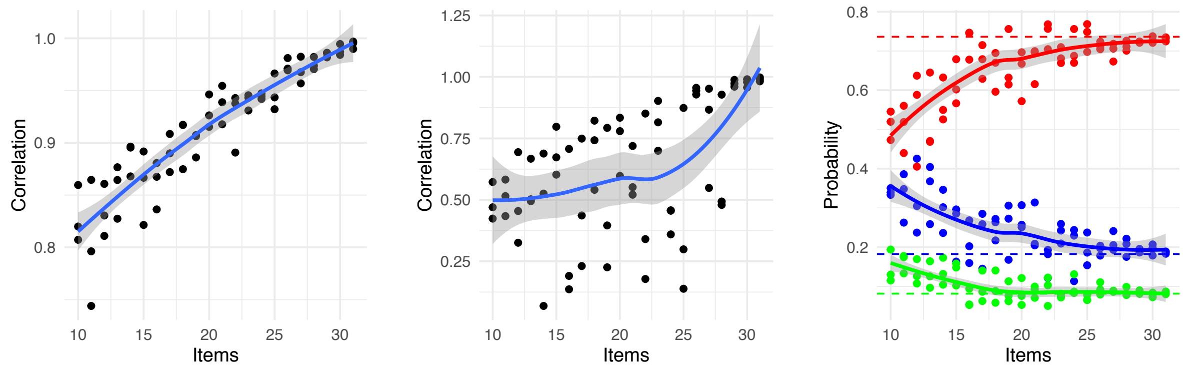
precise claims about which survey respondents have moderate ideal points, however defined. And we want to know when we can make precise claims about which respondents are very likely to be Downsians, as indicated by the size of the associated parameter. We also want to know how close our average estimates will be for all three parameters that indicate the fraction of the respondents that are Downsian, Conversian, and inattentive types.

Figure A1 shows the results for the first simulation method. The leftmost panel shows the correlation between the estimated ideal points in a given trial and the estimated ideal points using all 32 items on the y-axis. The x-axis is the number of used items in each trial. We fit a LOESS smoother to this relationship. With only 10 items this correlation hovers at a little over 0.8 on average but with correlations as low as .74. The relationship is close to linear. Twenty items are needed to consistently achieve correlations above .9, though of course more items are better.

The middle panel shows the correlation for the Downsian probabilities, w_1, estimated in each trial and the Downsian probabilities estimated with 32 items. This relationship is much noisier, but ranges from .5 in expectation with 10 items to very close to 1 with 32.

The last panel shows the averages for each set of probabilities along with a horizontal line for the averages when 32 item are used. Green indicates w_1, blue indicates w_2 and red indicates w_3. It is clear from this graph that these estimates are severely biased with only 10 items. w_1 and w_2 are biased upwards and w_3 is biased downwards. We suspect that this is part of a more general bias towards equality of the three probabilities in small samples. The bias ranges from almost .25 to close to 0, with the relationship flattening substantially around 20 items.

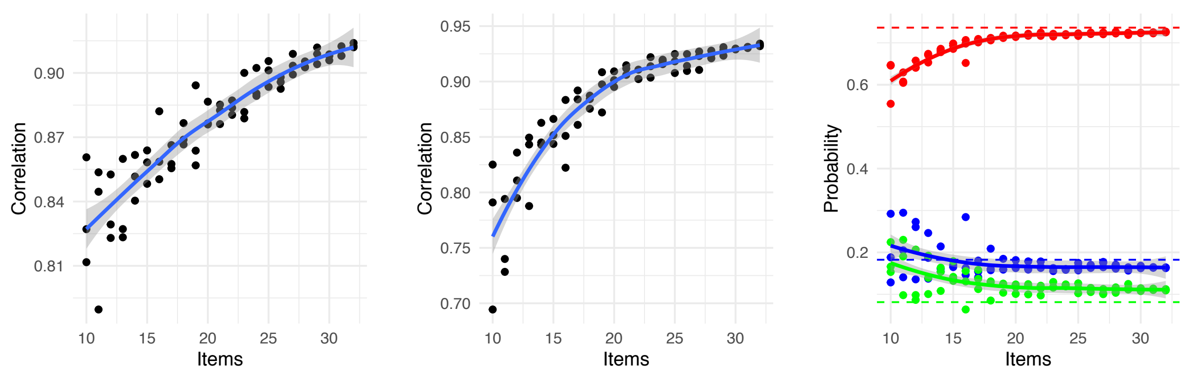
Figure A2 shows the results for the second simulation method. Assuming that our model is correctly specified substantially improves all of the metrics, particularly when few items are used. The association between the ideal points improves linearly in M from about .83 to about .91. The correlation in the probabilities improves rapidly from about .77 to about .94, flattening substantially around M=20. For the averages of the probabilities we see a



(a) Corr. ideal points (b) Corr. Pr($Downsian$) (c) Estimated composition

Figure A1: Results from Simulation 1

smaller but still substantial bias around M=10, which is mostly eliminated by M=20. In each case a small discrepancy remains between the benchmark parameters and the estimated parameters, reflecting a small degree of model misspecification.



(a) Corr. ideal points (b) Corr. Pr($Downsian$) (c) Estimated composition

Figure A2: Results from Simulation 2

There is no objective criteria for what threshold of items to use for precise estimation of these parameters. We consider the estimates using only 10 items to be clearly inadequate. It is clear from the graphs that greater numbers of items are better, and even greater than 32 would be preferable. However given the data available to us we choose to make do with datasets of 20 items or more. These estimates retain a small amount of bias against one of our central conclusions: that a low dimensional model is a good characterization of the preferences of most individuals. However they do not contain so much bias as to make type 2 errors very likely.

B.2 Other features of policy items that might improve model accuracy

So far we have only considered the number of items as an indicator of the power of a given dataset. However there are several other considerations that one might take into account in assessing power. The informativeness of a given dataset will depend on the unknown item parameters in complex ways. For instance, items that divide extreme liberals from moderate liberals are informative with respect to the parameters of those respondents but may not be very informative with respect to conservative respondents. So the position of the estimated cut points matter, and so does heterogeneity in these cut points. Items that are less discriminating will yield noisier estimates as well. In other words, "bad" items lead to "bad" estimates.

These factors are difficult to assess a priori. The margin of the survey question may be used as a rough indicator of where in the spectrum of ideal points the question is likely to be discriminating. In our case all survey questions are coded in what we believe to be the "conservative" direction. We can use the standard deviation of the margins as a measure of the coverage of these items. We evaluate whether the dispersion of the margins is an important factor in our simulated datasets, leaving a more thorough assessment of this methodological question to future work.

Table A1 shows the estimates from three models where the dependent variable is the correlation between the estimated w_1s from each simulation and the estimated w_1s using all 32 items from the CCES. These simulations are from our first method, described above. These models use three explanatory variables: the number of items, the standard deviation of the margins, and the interaction between those two factors. The number of items explains about a quarter of the variation in this correlation. However the standard deviation of the margins explains little if any variation, and only slightly improves upon a model using only the number of items.

Table A1: Effect of the number of items and standard deviation of the question margins on the correlation between the estimated w_1 and the benchmark w_1

	Dependent variable:		
	$corr_{w_1}$		
	(1)	(2)	(3)
# of items	0.021***		−0.054
	(0.004)		(0.050)
SD(margins) × # of items			0.560
			(0.371)
SD(margins)		1.832	−6.776
		(2.172)	(5.865)
Constant	0.204**	0.391	1.112
	(0.093)	(0.288)	(0.786)
Observations	66	66	66
R^2	0.265	0.011	0.299
Adjusted R^2	0.254	−0.004	0.266
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table A2 shows estimates using the same independent variable, but here the dependent variable is the correlation between the simulated ideal points and the benchmark ideal points. This time the number of items explains 86% of the variance in the correlation. The standard deviation of the margins adds little if any explanatory power.

We take these models as evidence that, at least in this dataset, the number of items is a much more important factor than having a lot of dispersion in the margins. This may be because any random sample of the items available is sufficiently dispersed. However for our purposes we opt for a simple inclusion criterion and use all datasets where respondents answer at least 20 questions.

Table A2: Effect of the number of items and standard deviation of the question margins on the correlation between the estimated ideal points and the benchmark ideal points

	Dependent variable:		
	$corr_x$		
	(1)	(2)	(3)
# of items	0.009***		0.010**
	(0.0004)		(0.005)
SD(margins) × # of items			−0.013
			(0.038)
SD(margins)		0.186	0.286
		(0.494)	(0.595)
Constant	0.740***	0.890***	0.702***
	(0.009)	(0.066)	(0.080)
Observations	66	66	66
R^2	0.859	0.002	0.859
Adjusted R^2	0.856	−0.013	0.853
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table A3 shows the median number of responses to policy items for 11 large-sample surveys of political views: the 2006-2016 Cooperative Congressional Election Studies and the 2000 and 2004 National Annenberg Election Surveys. The surveys where the median respondent answers at least 20 policy questions are the 2012, 2013, 2014, 2015, 2016, 2017 and 2018 Cooperative Congressional Election Studies, the data sets represented in the paper.

Table A3: Number of policy items on large sample surveys

Survey	Median Policy Responses
CCES 2006	12
CCES 2007	13
CCES 2008	14
CCES 2009	11
CCES 2010	16
CCES 2011	13
CCES 2012	21
CCES 2013	22
CCES 2014	32
CCES 2015	33
CCES 2016	28
CCES 2017	31
CCES 2018	35
NAES 2000	17
NAES 2004	9

C Model Validation with the Stanford Module of the 2010 CCES

In this appendix, we show that the example given in Figure 2 of the main text generalizes. If we compare any two questions out of the 133 question asked on the 2010 CCES module, respondents classified as Downsian moderates are more likely to give spatially consistent responses. Downsian moderates become even more likely to give spatially consistent responses when the magnitude of any inconsistency would be large. Conversians are more likely to give spatially inconsistent responses and their likelihood of doing so depends less on the magnitude of the inconsistency. This validation exercise requires no knowledge of our model to understand.

Consider a 133-by-133 matrix where each row and column represents one of our 133 items. The rows are ordered by support for the liberal alternative such that the top row is the least popular liberal policy and the bottom row the most popular liberal policy. The columns are ordered by support for the conservative policy. In this arrangement, the bottom left of our graph represents item pairs where the liberal alternative is very popular for the item in the rows and the conservative alternative is very popular for the item in the column. As we ascend towards the top right of the matrix, the liberal alternative becomes less and less popular for the row item, and the conservative alternative becomes less and less popular for the column item.

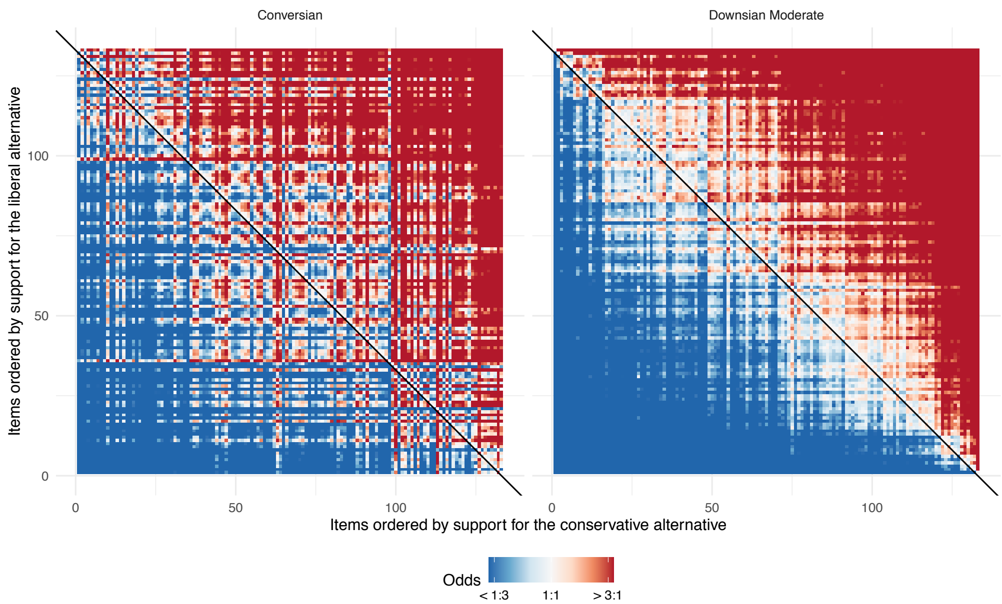
Without any statistical model, we want to try to capture the proportion of "spatial errors." If the items were perfectly Guttman scalable, then the margins would be sufficient. Consider a pair of items on the bottom left of our matrix, where the row item is R and the column item is C. Let 1 be a conservative response and 0 be a liberal response. Giving both liberal responses or both conservative responses will always be spatially consistent. For items on the bottom left, giving liberal responses to row items, $R = 0$, and conservative responses to column items, $C = 1$, is also spatially consistent, because these responses represent the majority of respondents. $R = 0$ and $C = 1$ is the moderate response to both questions. For the row item the conservative response is rare, and therefore relatively extreme, and for the column item the liberal response is rare, and therefore relatively extreme. So the response pattern $R = 1$, $C = 0$ gives the extreme conservative response on one question and the extreme liberal response on the other. This is a spatial error that suggests the respondent giving this answer pair has views not well summarized by a single dimension of policy ideology.

In the bottom left of the matrix, $(R = 0, C = 1)$ represents a spatially consistent choice and $(R = 1, C = 0)$ also represents a spatially inconsistent choice. As we move up the rows and to the left on the columns the margins of the questions get closer. At some point, the situation flips. Once majorities support the conservative side on the rows and the liberal side on the columns, then $(R = 1, C = 0)$ represents a spatially consistent choice and $(R = 0, C = 1)$ represents a spatially inconsistent choice. If these choices were perfectly Guttman scalable, than we would no longer observe the spatially inconsistent choice. In a random utility model, errors should become more common as the margins of the question become closer.

Figure A3 graphs the odds of choosing $(R = 1, C = 0)$ against $(R = 0, C = 1)$ for respondents who are classified as Conversian (left frame) and Downsian (right frame) moderates by our model. Moderate here indicates someone whose ideal point is in the middle third of the distribution with higher posterior probability Downsian than Conversian or inattentive. We focus on moderates to support the claims we make about moderates specifically in the paper.

Our expectation is that, for subjects whose views are well-explained by a single dimension of ideology, the odds should should be low on the bottom left, when $(R = 1, C = 0)$ is the spatially inconsistent choice, and high on the top right, when $(R = 1, C = 0)$ is the spatially consistent choice. The odds should approach 1:1 in the middle when the most spatially consistent choice is to respond in the same direction to both questions.

Figure A3: Odds of Spatially Consistent versus Spatially Inconsistent Choices



For each pair of 133 issues on the 2010 CCES, the color of each "pixel" represents the odds of a randomly selected respondent giving the conservative answer to the question indicated by the pixel's x-axis position and the liberal answer to the question indicated by the pixel's y-axis position from among those respondents giving one conservative and one liberal answer to that question pair. The questions are ordered by support for the conservative position on the x-axis and by support for the liberal position on the y-axis.

Under perfect one-dimensional spatial voting, the data would be Guttman scalable. In that case, these odds would be greater than 1:1 everywhere above the -45 degree line and less than 1:1 everywhere below the -45 degree line. Notice that for those respondents who we identify as Downsian moderate this is largely the case. On the other hand, for those respondents identified as Conversian, there is a great deal of red (odds greater than 1:1) below the -45 degree line and blue (odds less than 1:1) above the -45 degree line. It is clear from the graph that Conversians are much less constrained than are Downsian moderates.

This analysis provides descriptive nonparametric evidence that our model successfully separates ideologically consistent moderates (Downsians) from those whose responses are much less constrained by the ideological dimension (Conversians).

D Modeling Spatial Preferences in Two Dimensions

In the model presented in the main text, voters can either hold one-dimensional spatial preferences (with error) or hold issue opinions that are (across all such voters) independent across issues (Conversians and inattentives). An alternative approach would be to place all voters in a higher-dimensional preference space. Indeed, putting aside the small number of inattentive voters, it can be easily demonstrated that the mixture model that we advance can be represented as, and is isomorphic to, a standard two-dimensional model in which our Downsians have ideal points that fall on a single line and the Conversians fall on a single point that lies away from that line.[2] To see this, recall that, in the notation introduced in the main text, the probability that a Downsian respondent i answers issue question j in the affirmative ($y_{ij} = 1$) is

$$\Lambda\left(\beta_j(x_i - \alpha_j)\right).$$

If we extend this spatial choice function to two dimensions, the probability that $y_{ij} = 1$ becomes

$$\Lambda(\tilde{\alpha}_j + \tilde{\beta}_{j1}\tilde{x}_{i1} + \tilde{\beta}_{j2}\tilde{x}_{i2}).$$

While adding a second dimension to the usual quadratic spatial preference model increases the number of x and β parameters that characterize each choice, there is still only one α parameter (Clinton, Jackman, and Rivers, 2004, p. 365). The definition of $\tilde{\alpha}$ differs from its one-dimensional counterpart which is why we place a tilde over it (and the other parameters in the two-dimensional model). Note that $\tilde{\alpha}_j = -\alpha_j\beta_j$ in the one-dimensional case (in which $\tilde{\beta}_{j2} = 0$ for all j). Now, suppose that the data are generated according to the mixture model presented in the main text. We can represent the choice probabilities of Downsians in that model by setting $\tilde{\alpha}_j = -\alpha_j\beta_j$, $\tilde{\beta}_{j1} = \beta_j$, $\tilde{x}_{i1} = x_i$, and $\tilde{x}_{i2} = 0$ for all (Downsian) respondents

[2] We thank Ben Lauderdale for first pointing this out to us.

i and issue questions j because then

$$\Lambda(\tilde{\alpha}_j + \tilde{\beta}_{j1} x_i + \tilde{\beta}_{j2} \cdot 0)$$

equals

$$\Lambda\left(\beta_j(x_i - \alpha_j)\right).$$

Holding fixed these values of $\tilde{\alpha}$ and $\tilde{\beta}_{j1}$, we can accommodate the Conversian voters by setting their $\tilde{x}_{i1} = 0$ and their $\tilde{x}_{i2} = 1$ and choosing $\tilde{\beta}_{j2}$ to solve

$$\lambda_j = \Lambda(\tilde{\alpha}_j + \beta_j \cdot 0 + \tilde{\beta}_{j2} \cdot 1)$$

for all (Conversian) i and j. Rearranging we have

$$\Lambda^{-1}(\lambda_j) = \tilde{\alpha}_j + \tilde{\beta}_{j2} \cdot 1$$

or

$$\tilde{\beta}_{j2} = \Lambda^{-1}(\lambda_j) - \tilde{\alpha}_j.$$

Adding the inattentive voter type to the mix breaks the isomorphism of the two models, but given that few respondents of this type are estimated to exist in the data, the two models are close to isomorphic in this application.[3] Because spatial models in two dimensions are invariant to translations, dilations, reflections, and rotations of the ideal point space (see Clinton, Jackman, and Rivers, 2004, p. 365–366), there is a continuum of ways in which the model presented in the main text (leaving out the inattentives) can be made isomorphic to a (restricted) two-dimensional spatial model. However, all of these isomorphic two-dimensional models have the Downsians falling on a single line through the two-dimensional space and

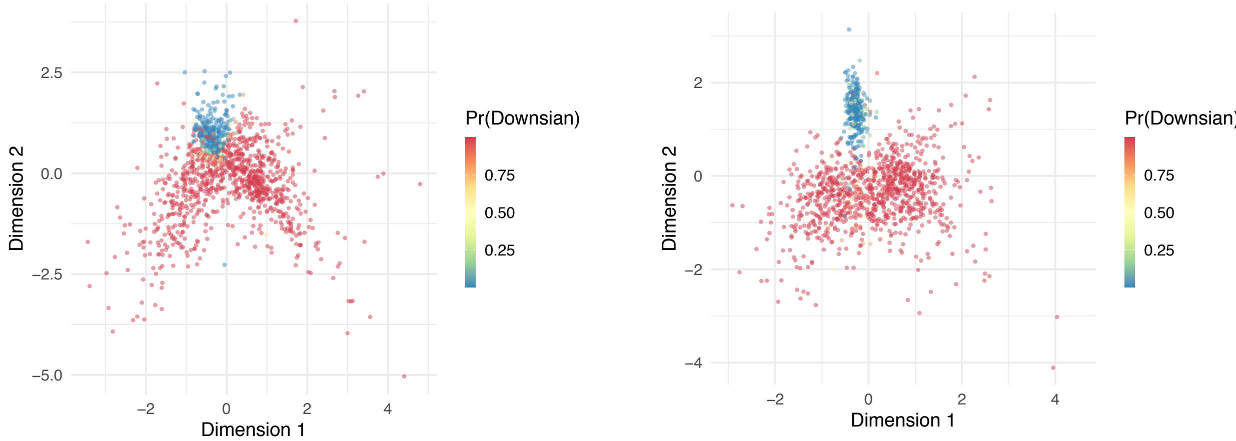
[3] Adding a third spatial dimension would be sufficient to recreate the isomorphism with inattentives included.

the Conversians falling on a point that does not (in general) lie on that line.[4]

In this Appendix, we allow for the possibility that (some) voters have two-dimensional spatial preferences. We focus this exploration on the same 133-question dataset drawn from the 2010 CCES that we employ in Appendix C, the 2010 CCES module dataset. The large number of issue items found in this dataset relative to the other datasets presented in the text gives us the best opportunity to explore preferences in more than one dimension. We also present estimates of the out-of-sample fit of various alternative preference models considered for all of the datasets analyzed in the text.

We first apply a standard two-dimensional IRT-like model (Clinton, Jackman, and Rivers, 2004) to the 2010 CCES module dataset. Panel (a) of Figure A4 plots the resulting estimated ideal points. The points are colored according to the estimated probability that a respondent is a Downsian as estimated by the mixture model employed in the text. This plot does not reveal a single line of Downsians and a single point of Conversians that falls away from that line. However, the deviation from that pattern is perhaps less stark than it might appear. First, we see that the Conversians are concentrated in a small area of the graph. Second, because there is a stochastic component to the voters' preferences and because their locations are determined by no more than 133 questions (91.9 on average), each ideal point is estimated with error. Therefore, even if the true ideal points all fell on a single line in the space, we would expect the estimates to form a cloud around that line. To demonstrate this, Panel (b) of Figure A4 shows the estimated results when the same two-dimensional spatial model is applied to a simulated data set produced according to our mixture model calibrated to the CCES 2010 module data. Here we see that despite the mixture model holding exactly in the data, the Conversians are clustered, but do not fall on a single point nor do the

[4]If, in the parameterization presented, $\Lambda^{-1}(\lambda_j) = \tilde{\alpha}_j$ for all j then $\tilde{\beta}_{j2} = 0$ for all j and Conversians would be located at $\tilde{x}_i = (0, 0)$ which is a point on the line containing the Downsians. Of course, in this case Conversians cannot be empirically distinguished from Downsians because their choice probabilities would be identical to those of Downsians for whom $x_i = 0$. Note that in this knife-edged case where there is only one-dimension of choice, the values of $\tilde{\beta}_{j2}$ and $\tilde{x}_{i2}$ are not separately identified because $\tilde{\beta}_{j2} = 0$ for all j with $\tilde{x}_{i2} \in (-\infty, \infty)$ for all i and $\tilde{x}_{i2} = 0$ for all i with $\tilde{\beta}_{j2} \in (-\infty, \infty)$ for all j yield equivalent choice probabilities.



(a) Estimated (b) Simulated

Figure A4: Estimating respondent preferences in two spatial dimensions. *Panel (a) shows the locations of 2010 CCES module respondents as estimated by a standard two-dimensional spatial model. The points are shaded to reflect the probability that each respondent is of the Downsian type as estimated by the model presented in the main text. Panel (b) shows the same plot based on simulated data that is calibrated to the 2010 CCES module dataset under the assumptions of the model presented in the main text.*

estimated locations of the Downsians fall on a single line. The general pattern shown in the two panels is similar though the locations of the Conversians is more strongly differentiated in the simulated data and there appears to be more structure to the second dimension in the empirical data. Given that there is no second dimension of spatial preference in the simulated data, this is not surprising. Though there is apparent structure in the second dimension of the empirical data, the first and second dimension locations are far from independent calling into question the degree to which there is an important distinct second dimension of preference manifest in the issue question responses.

Indeed, the empirical estimates reveal the horseshoe pattern often found when two-dimensional scaling models are applied in situations in which a single underlying dimension is expected (see Diaconis, Goel, and Holmes, 2008). In such cases, the recovered second dimension can be accounting for some misspecification of the functional form of the stochastic spatial preference, choice or distance function rather than a distinct second dimension (for example, in our context, distinct "economic" and "social" policy preference dimensions)

(Kendall, 1970; Shepard, 1974; Hill and Gauch, 1980; Diaconis, Goel, and Holmes, 2008; de Leeuw, 2011*a*).

Because there may be a distinct second dimension of spatial preference or the assumed functional form of the one-dimensional spatial preference model may be driving our results, we next consider how the inclusion of a second dimension into the mixture model affects our estimates of the fraction of Downsians and Conversians in the population. To do this, we fit an extended version of our mixture model to the CCES 2010 module dataset that allows the Downsians to have preferences over two spatial dimensions.

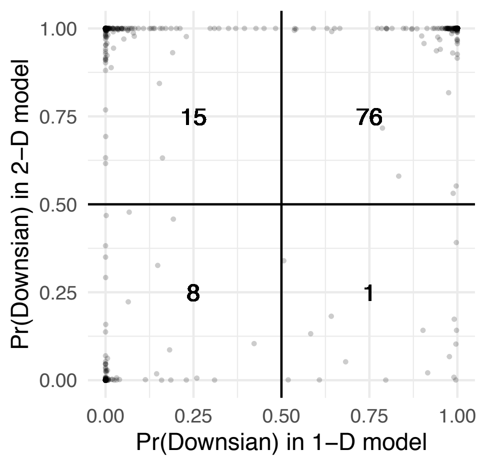


Figure A5: Estimated probabilities of each respondent being of the Downsian type.

For each respondent in the 2010 CCES module data, the x-axis shows the probability that a given respondent is of the Downsian type when one-dimensional spatial preferences are assumed. The y-axis shows the probability that a given respondent is of the Downsian type when two-dimensional spatial preferences are assumed. The quadrants partition respondents predicted to be Downsian from those predicted to be non-Downsian in either model or both. The numbers indicate the percentage of the sample that is estimated to fall into each quadrant. For example, 76 percent of the sample is estimated to be of the Downsian type in both one and two dimensions, while one percent of the sample is estimated to be Downsian when one spatial dimension is assumed, but non-Downsian when two spatial dimensions are assumed.

Figure A5 shows the estimated probability of being a Downsian for each survey respondent under the one-dimensional and two-dimensional mixture models. The four quadrants of the plot contain voters who are estimated to be (moving clockwise from the upper left): Downsian in the two-dimensional model, but Conversian in the one-dimensional model; Downsian in both models; Downsian in the one-dimensional model and Conversian in the two-dimensional model; and Conversian in both models. Whereas the one-dimensional model estimates about 23 percent of the sample to be Conversian, the two-dimensional model places only 9 percent of the sample in that category. Fifteen percent of the sample moves from Conversian to Downsian when a second dimension is available whereas only one percent moves from Downsian to Conversian. As noted in the main text, this suggests that some of the voters identified as Conversian moderates in the main text may hold preferences that, while not easily reconciled with a single spatial dimension, can be made reconcilable with spatial preferences when a second spatial dimension is added. Thus, our characterization of the fraction of "moderates" who actually have spatial preferences is perhaps understated.

Another related question is whether the addition of a second spatial dimension substantially improves the fidelity of the model with the data. To answer that question, we need a measure of (out of sample) fit. Table A4 reports the in-sample log likelihood as well as the out-of-sample perplexity associated with each model when applied to the 2010 CCES module dataset. The out-of-sample perplexity is approximated via a five-fold cross validation. Under the "null" model, across the entire sample, preferences are assumed to be independent across choices (in effect, all voters are assumed to be Conversians). The "1-D (mix.)" is the model presented in the main text that considers a mixture of Downsian, Conversian, and inattentive respondents. The "2-D (no mix.)" is the standard two-dimensional model used to produce the estimates in Figure A4. It does not include Conversian and inattentive types. The "2-D (Mix.)" is a version of the model used in the main text in which Downsians are given preferences over two spatial dimensions rather than one, and includes Conversian and inattentive types. Given the large number of data points (1,300 respondents answering on average 91.9 issue questions), it is not surprising that the differences between each pair of log likelihoods are statistically significant (p values not shown). That is, a statistically significant increase in data fit is afforded by each increase in model complexity.

Model	Log-likelihood	Perplexity
Null	-72675	1.84
1-D (mix.)	-53049	1.58
2-D (no mix.)	-51978	1.57
2-D (mix.)	-51383	1.56

Table A4: Model log likelihood and perplexity, 2010 CCES module dataset. *Shows the estimated model log likelihoods and estimated average (per item) perplexities across four possible models of preference. Each model is fit to the same 1,300 respondents answering an average of 91.9 issue questions). Each row of the table presents the estimated fit for a given model. The rows are organized in increasing order of model complexity. The log likelihood is estimated in sample. Perplexity is estimated out of sample using five-fold cross validation. The differences in log likelihood are highly statistically significant though the reductions in perplexity as model complexity increases are modest (except when comparing the null model to the others). Each model is described in the text.*

However, the perplexity differences among the various spatial models are modest particularly in comparison to the null model. Perplexity can be understood as the average number of bits per issue item required to compactly represent the responses of a single respondent. The higher the likelihood the model assigns to each observed pattern of the data the lower the perplexity (the perplexity is the average of the inverse of the geometric mean probability of the responses given by each respondent). If every respondent were an inattentive type, perplexity would be 2, which is the theoretical maximum (the maximally entropic data generating process). On the other hand, if every respondent expressed one of only two patterns across items, the perplexity would approach 0 (1 over the number of items) because a single bit would be sufficient to label the two observed patterns. As with the log likelihood, the value of perplexity is a function of both the nature of the data and the fidelity of the model. Because the perplexity is calculated using cross-validation, the observed reduction in the estimated perplexity as model complexity increases is not a mechanical result.

In fact, only small improvements in model fit result from the addition of a second spatial dimension. The inclusion of the Conversian and inattentive types appears to increase the fit of the two-dimensional model. However, the differences in fit among the various models that include a spatial component are very small (less than 1 percent differences in perplexity

		Log likelihood				Perplexity			
				2-D				2-D	
Survey	Avg. no. of items	Null	1-D Mix.	No mix.	Mix.	Null	1-D Mix.	No mix.	Mix.
2012	18.6	-661043	-550889	-551832	-550929	1.93	1.74	1.75	1.74
2013	21.8	-226870	-194338	-193809	-193867	1.89	1.74	1.74	1.74
2014	31.6	-1156518	-960685	-956933	-949695	1.92	1.74	1.73	1.72
2015	32.5	-292295	-230744	-227953	-226884	1.89	1.66	1.66	1.65
2016	28.8	-1137878	-972217	-973954	-971166	1.86	1.71	1.71	1.71
2017	30.9	-358915	-269839	-269580	-267369	1.90	1.64	1.64	1.63
2018	33.1	-1274454	-973250	-975855	-968311	1.91	1.66	1.67	1.66

Table A5: Model log likelihood and perplexity, 2012–2018 CCES datasets. *Shows the estimated model log likelihoods and estimated average (per item) perplexities across four possible models of issue preference. Each model is fit to the same respondents to each survey. The average number of responses to each survey is given in the table. Each row of the table presents estimated model fits for a given survey. The differences in log likelihood are statistically significant across the models for each survey though the reductions in perplexity as model complexity increases are very small (except when comparing the null model to the others). Each model is described in the text.*

per issue item). Table A5 shows the log likelihoods and perplexities associated with the Null, 1-D (with mixture), 2-D (without mixture), and 2-D (with mixture) models described above when applied to the CCES datasets from 2012 to 2018 analyzed in the text. As with the 2010 CCES module dataset, adding model complexity increases fit in a statistically significant way (the log likelihoods differ by more than chance would allow). However the degree of additional (out of sample) fit is minimal (often zero to two decimal places).

E Additional Results on Selection and Accountability

Table 4 in the text assessed the extent to which the voting behavior of different types of individuals responds to candidate ideology and experience. To assess the extent to which each group contributes to election results, we utilized a trichotomous dependent variable that takes a value of 1 if the respondent voted for the Democratic candidate, 0 if the respondent voted for the Republican candidate, and 0.5 if the respondent abstained or voted for a third-party candidate.

For readers interested in the extent to which those previous results were explained by voter turnout versus vote choice, we replicate those analyses but utilize alternative dependent variables. Table A6 excludes those who abstained or supported a third-party candidate and utilizes a binary dependent variable indicating support for the Democratic candidate. This analysis suffers from the potential concern that the independent variables of interest could affect turnout, which could induce bias. However, if we assume that candidate ideology and experience do not influence turnout, we can interpret these results as the differential effects of ideology and experience for those who voted.

If anything, the interactive coefficients in Table A6 are greater than those in Table 4. In other words, if we condition on those who voted, moderate, Conversian, and inattentive individuals are even more likely than liberals and conservatives to change their partisan vote choices in response to candidate ideology and experience. Of course, moderate, Conversian, and especially inattentive individuals are less likely to vote than liberal and conservative individuals, so these estimates overstate the extent to which these groups contribute to election results. But these results show that among those who vote, the non-ideologues are especially likely to contribute to electoral selection and accountability.

Additionally, Table A7 shows the same analyses but utilizes abstention as the dependent variable of interest. Consistent with our previous results, we find that moderate, Conversian, and inattentive individuals are more likely to abstain than liberals and conservatives.

The first three columns show that the extent to which these groups differentially abstain does not meaningfully vary as the ideologies of the candidates shift from favoring the Republican candidate to favoring the Democratic candidate.

However, we do find that the participation differences do vary across candidate experience in ways that we might expect. As the experience gap between the Democratic and Republican candidate increases, conservatives become much more likely to abstain relative

Table A6: Excluding Abstainers

	DV = House Vote (Dem = 1, Rep = 0)								
	X = Ideological Midpoint			X = Incumbency			X = Experience		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
X*Moderate	.107	.102	.101	.219	.214	.210	.225	.221	.215
	(.021)	(.020)	(.020)	(.013)	(.013)	(.013)	(.013)	(.013)	(.013)
X*Conversian	.157	.146	.150	.253	.242	.233	.257	.248	.237
	(.024)	(.024)	(.024)	(.015)	(.015)	(.015)	(.016)	(.016)	(.015)
X*Inattentive	.106	.096	.098	.250	.229	.221	.251	.228	.219
	(.051)	(.046)	(.046)	(.031)	(.029)	(.029)	(.031)	(.029)	(.029)
X*Conservative	.011	.007	.010	.001	.014	.016	.002	.018	.017
	(.012)	(.013)	(.014)	(.008)	(.010)	(.011)	(.009)	(.010)	(.011)
X	.042	-.012		.066	-.039		.065	-.059	
	(.009)	(.012)		(.005)	(.013)		(.005)	(.012)	
Moderate	-.492	-.480	-.479	-.513	-.505	-.498	-.523	-.515	-.506
	(.011)	(.011)	(.011)	(.008)	(.008)	(.008)	(.008)	(.008)	(.008)
Conversian	-.460	-.448	-.451	-.474	-.467	-.462	-.484	-.477	-.471
	(.014)	(.014)	(.014)	(.010)	(.009)	(.009)	(.010)	(.010)	(.010)
Inattentive	-.520	-.509	-.511	-.558	-.549	-.543	-.563	-.553	-.547
	(.029)	(.026)	(.026)	(.019)	(.018)	(.018)	(.019)	(.018)	(.018)
Conservative	-.917	-.894	-.892	-.890	-.878	-.866	-.891	-.879	-.866
	(.006)	(.007)	(.007)	(.005)	(.005)	(.006)	(.005)	(.006)	(.006)
Year FEs	✓	✓		✓	✓		✓	✓	
District FEs		✓			✓			✓	
District-Year FEs			✓			✓			✓
Observations	102,350	102,350	102,350	143,715	143,715	143,715	143,715	143,715	143,715

District-clustered standard errors in parentheses. Liberals are the omitted category.

to liberals, and moderates, Conversians, and inattentive individuals are somewhere in between. In other words, an experience advantage for the Republican (Democratic) candidate motivates conservative (liberal) individuals to participate relative to liberal (conservative) individuals. Interestingly, the estimated differences between conservatives and moderates are greater than those between liberals and moderates. One potential explanation is that conservative abstention is more responsive to candidate experience than liberal abstention or that moderate abstention more closely matches that of liberals.

Table A8 replicates the analyses in Table 4 but adds in controls for party identification. Specifically, all regressions include fixed effects for each possible category of the seven-point party identification scale. On one hand, these controls might increase precision since party identifcation is strongly correlated with vote choice. On the other hand, controlling for party identification could induce bias because the ideology and experiences of congressional candi-

Table A7: Analyzing Abstention

	DV = House Abstention (Abstain/Other = 1, Dem/Rep = 0)								
	X = Ideological Midpoint			X = Incumbency			X = Experience		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
X*Moderate	-.026	-.033	-.033	.028	.030	.031	.031	.035	.035
	(.016)	(.016)	(.016)	(.012)	(.012)	(.012)	(.012)	(.012)	(.012)
X*Conversian	-.011	-.014	-.010	.040	.042	.038	.041	.044	.043
	(.020)	(.019)	(.019)	(.012)	(.012)	(.012)	(.013)	(.012)	(.012)
X*Inattentive	-.009	-.025	-.019	.044	.042	.039	.071	.069	.068
	(.029)	(.029)	(.029)	(.019)	(.019)	(.019)	(.020)	(.020)	(.020)
X*Conservative	-.008	-.015	-.014	.118	.133	.129	.125	.144	.140
	(.017)	(.017)	(.017)	(.016)	(.015)	(.015)	(.016)	(.016)	(.016)
X	.036	.017		-.051	-.056		-.060	-.065	
	(.014)	(.015)		(.011)	(.016)		(.012)	(.017)	
Moderate	.242	.236	.234	.205	.196	.194	.203	.193	.191
	(.009)	(.009)	(.009)	(.008)	(.008)	(.008)	(.008)	(.008)	(.008)
Conversian	.263	.250	.243	.225	.208	.204	.223	.205	.201
	(.011)	(.010)	(.010)	(.008)	(.008)	(.008)	(.009)	(.009)	(.009)
Inattentive	.353	.345	.341	.319	.302	.300	.304	.287	.284
	(.017)	(.017)	(.018)	(.013)	(.013)	(.013)	(.014)	(.014)	(.014)
Conservative	-.049	-.056	-.055	-.096	-.107	-.104	-.103	-.116	-.113
	(.009)	(.009)	(.009)	(.008)	(.008)	(.008)	(.009)	(.008)	(.008)
Year FEs	✓	✓		✓	✓		✓	✓	
District FEs		✓			✓			✓	
District-Year FEs			✓			✓			✓
Observations	159,006	159,006	159,006	233,445	233,445	233,445	233,445	233,445	233,445

District-clustered standard errors in parentheses. Liberals are the omitted category.

dates could potentially influence the reported party identification of respondents. Because of this potential bias, we believe analyses that exclude partisanship controls are more reliable.

When we control for party identification, the estimated interactive effects of interest in Table A8 are similar to those in Table 4 although slightly attenuated. This could follow from the relative appeal of Democratic and Republican congressional candidates affecting reports of party identification. Nevertheless, even when we control for party, the results are qualitatively similar.

We might also want to know how the ideological types we identify interact with party identification. Because party identification is strongly correlated with vote choice, we would expect, for example, liberal Democrats to behave differently than liberal independents. To assess this possibility, we coded indicators for every potential combination of our ideological types (liberal, moderate, conservative, Conversian, and inattentive) and three-point party

Table A8: Controlling for Party ID

	DV = House Vote (Dem = 1, Rep = 0, Abstain/Other = .5)								
	X = Ideological Midpoint			X = Incumbency			X = Experience		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
X*Moderate	.036	.037	.039	.044	.046	.045	.043	.045	.044
	(.010)	(.010)	(.010)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)
X*Conversian	.040	.043	.046	.048	.049	.046	.044	.045	.042
	(.012)	(.012)	(.011)	(.007)	(.007)	(.007)	(.008)	(.008)	(.008)
X*Inattentive	.015	.016	.018	.022	.020	.013	.016	.014	.005
	(.017)	(.017)	(.017)	(.012)	(.012)	(.012)	(.012)	(.012)	(.012)
X*Conservative	.024	.024	.026	.015	.018	.017	.007	.010	.009
	(.013)	(.013)	(.014)	(.011)	(.011)	(.012)	(.012)	(.012)	(.012)
X	.017	-.007		.057	.015		.061	.005	
	(.008)	(.008)		(.007)	(.010)		(.007)	(.010)	
Moderate	-.188	-.185	-.186	-.183	-.181	-.180	-.183	-.181	-.180
	(.006)	(.006)	(.006)	(.005)	(.005)	(.005)	(.005)	(.005)	(.005)
Conversian	-.176	-.173	-.176	-.168	-.165	-.166	-.167	-.164	-.165
	(.007)	(.007)	(.006)	(.005)	(.005)	(.005)	(.005)	(.005)	(.005)
Inattentive	-.184	-.180	-.183	-.180	-.177	-.176	-.178	-.174	-.172
	(.010)	(.009)	(.009)	(.008)	(.008)	(.008)	(.008)	(.008)	(.008)
Conservative	-.371	-.367	-.368	-.350	-.348	-.348	-.347	-.345	-.345
	(.007)	(.007)	(.007)	(.006)	(.006)	(.006)	(.007)	(.007)	(.007)
Year FEs	✓	✓		✓	✓		✓	✓	
District FEs		✓			✓			✓	
District-Year FEs			✓			✓			✓
Party ID FEs	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	152,616	152,616	152,616	224,047	224,047	224,047	224,047	224,047	224,047

District-clustered standard errors in parentheses. Liberals are the omitted category.

identification (Democrat, independent, and Republican). We then replicated the methodology used in Table 4 but separately examined each of these categories. The results of this analysis are in Table A9.

As expected, both our ideological classifications and party identification are important for explaining voting behavior and the contributions of different voters to selection and accountability, and there are interesting interactions between ideology and party identification.

Among liberals, Republicans (a very small share of liberals) are more responsive to candidate ideology and experience than Democrats. Conversely, among conservatives, Democrats are more responsive than Republicans. Similarly, among Democrats, conservatives are more responsive than liberals, and among Republicans, liberals are more responsive than conservatives. These results are consistent with the possibility that party identification is another proxy for ideology. For example, liberal Republicans are likely more ideologically moderate than liberal Democrats, and since more ideologically moderate individuals are likely more responsive to candidate ideology and experience, we find that the former group is more responsive.

Interestingly, among moderates, independents are not necessarily more responsive to candidate ideology and experience than partisans. Moderate Democrats and moderate Republicans are among the most responsive groups. Similarly, Conversian Republicans are also very responsive to candidate ideology and experience.

The results in Table A9 suggest that if you want to understand the extent to which different people contribute to electoral selection and accountability, their ideological classification are more informative than their party identification. To be sure, independents are generally more responsive than partisans, but moderates and Conversians are much more responsive than liberals and conservatives. Furthermore, moderate and Conversian partisans appear to be more responsive than independent liberals and conservatives.

Table A9: Ideological Type by Party Identification

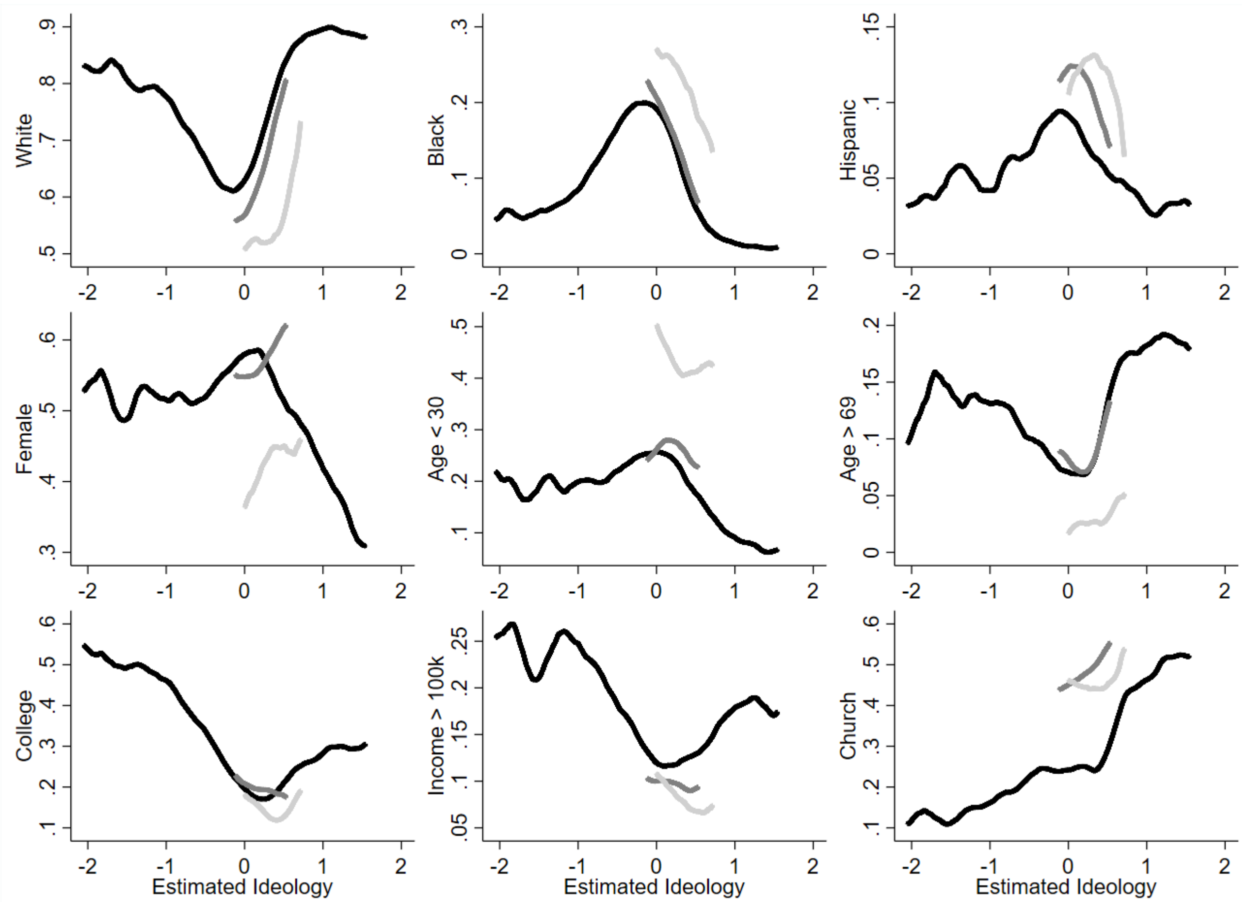
	DV = House Vote (Dem = 1, Rep = 0, Abstain/Other = .5)								
	X = Ideological Midpoint			X = Incumbency			X = Experience		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
X*Liberal Independent	.013	.012	.014	-.004	-.002	-.000	-.001	.001	.002
	(.013)	(.013)	(.013)	(.009)	(.009)	(.009)	(.009)	(.009)	(.009)
X*Liberal Republican	.030	.030	.021	.084	.090	.082	.105	.108	.096
	(.046)	(.045)	(.048)	(.036)	(.036)	(.036)	(.037)	(.037)	(.037)
X*Moderate Democrat	.048	.051	.055	.052	.051	.052	.055	.056	.057
	(.015)	(.015)	(.015)	(.010)	(.010)	(.009)	(.009)	(.009)	(.009)
X*Moderate Independent	.029	.028	.029	.024	.028	.026	.020	.023	.021
	(.013)	(.013)	(.013)	(.009)	(.009)	(.010)	(.010)	(.010)	(.010)
X*Moderate Republican	.051	.051	.051	.059	.066	.065	.058	.065	.064
	(.020)	(.020)	(.020)	(.014)	(.014)	(.014)	(.014)	(.014)	(.014)
X*Conservative Democrat	.062	.067	.069	.043	.046	.040	.066	.070	.067
	(.070)	(.065)	(.066)	(.035)	(.034)	(.034)	(.035)	(.034)	(.033)
X*Conservative Independent	.046	.045	.045	.027	.029	.027	.025	.026	.023
	(.018)	(.018)	(.019)	(.014)	(.014)	(.014)	(.014)	(.014)	(.014)
X*Conservative Republican	.017	.016	.020	.008	.014	.015	-.000	.006	.007
	(.015)	(.015)	(.016)	(.012)	(.013)	(.013)	(.013)	(.013)	(.013)
X*Conversian Democrat	.020	.021	.020	.038	.038	.034	.040	.039	.035
	(.016)	(.015)	(.015)	(.010)	(.009)	(.009)	(.010)	(.010)	(.010)
X*Conversian Independent	.063	.062	.065	.040	.040	.037	.035	.035	.033
	(.018)	(.017)	(.017)	(.010)	(.010)	(.010)	(.010)	(.010)	(.010)
X*Conversian Republican	.071	.075	.084	.070	.076	.075	.064	.072	.071
	(.020)	(.020)	(.020)	(.015)	(.015)	(.016)	(.015)	(.016)	(.016)
X*Inattentive Democrat	.007	.014	.023	.017	.016	.008	.013	.010	.001
	(.031)	(.030)	(.030)	(.019)	(.019)	(.019)	(.020)	(.019)	(.019)
X*Inattentive Independent	.061	.056	.050	.019	.019	.010	.007	.007	-.004
	(.022)	(.022)	(.021)	(.015)	(.015)	(.015)	(.015)	(.015)	(.015)
X*Inattentive Conservative	.003	.000	.007	.052	.050	.046	.054	.053	.048
	(.038)	(.037)	(.037)	(.025)	(.025)	(.025)	(.026)	(.026)	(.026)
X	.013	-.011		.060	.013		.061	.003	
	(.009)	(.010)		(.007)	(.011)		(.007)	(.011)	
Liberal Independent	-.096	-.094	-.095	-.081	-.084	-.084	-.083	-.085	-.085
	(.008)	(.008)	(.008)	(.007)	(.006)	(.006)	(.007)	(.007)	(.007)
Liberal Republican	-.340	-.334	-.328	-.355	-.353	-.350	-.366	-.363	-.358
	(.028)	(.028)	(.029)	(.023)	(.023)	(.023)	(.025)	(.024)	(.025)
Moderate Democrat	-.185	-.183	-.185	-.182	-.179	-.178	-.185	-.182	-.182
	(.009)	(.009)	(.009)	(.007)	(.007)	(.006)	(.007)	(.007)	(.006)
Moderate Independent	-.384	-.378	-.378	-.354	-.353	-.352	-.352	-.351	-.350
	(.008)	(.008)	(.008)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)
Moderate Republican	-.640	-.632	-.630	-.599	-.598	-.596	-.601	-.599	-.596
	(.011)	(.011)	(.011)	(.009)	(.009)	(.008)	(.009)	(.009)	(.009)
Conservative Democrat	-.455	-.448	-.450	-.420	-.417	-.411	-.431	-.428	-.424
	(.032)	(.031)	(.031)	(.023)	(.022)	(.021)	(.023)	(.022)	(.022)
Conservative Independent	-.724	-.717	-.716	-.666	-.665	-.664	-.667	-.664	-.663
	(.010)	(.010)	(.010)	(.008)	(.008)	(.008)	(.009)	(.009)	(.009)
Conservative Republican	-.771	-.762	-.763	-.720	-.718	-.717	-.717	-.715	-.714
	(.009)	(.009)	(.009)	(.008)	(.008)	(.008)	(.008)	(.009)	(.009)
Conversian Democrat	-.160	-.158	-.159	-.159	-.157	-.157	-.161	-.158	-.158
	(.010)	(.010)	(.010)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)
Conversian Independent	-.375	-.369	-.371	-.339	-.337	-.336	-.338	-.335	-.335
	(.009)	(.009)	(.009)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)
Conversian Republican	-.630	-.623	-.627	-.580	-.578	-.578	-.580	-.579	-.578
	(.011)	(.011)	(.011)	(.009)	(.009)	(.009)	(.009)	(.010)	(.010)
Inattentive Democrat	-.201	-.201	-.205	-.201	-.200	-.200	-.200	-.197	-.196
	(.020)	(.019)	(.019)	(.014)	(.014)	(.014)	(.015)	(.015)	(.014)
Inattentive Independent	-.402	-.393	-.391	-.361	-.358	-.356	-.356	-.353	-.349
	(.012)	(.012)	(.012)	(.010)	(.010)	(.010)	(.010)	(.010)	(.010)
Inattentive Conservative	-.584	-.574	-.579	-.557	-.551	-.551	-.559	-.553	-.553
	(.019)	(.019)	(.019)	(.014)	(.014)	(.014)	(.014)	(.015)	(.015)
Year FEs	✓	✓		✓	✓		✓	✓	
District FEs		✓			✓			✓	
District-Year FEs			✓			✓			✓
Observations	159,006	159,006	159,006	233,445	233,445	233,445	233,445	233,445	233,445

District-clustered standard errors in parentheses. Liberal Democrats are the omitted category.

F Demographics of Ideological Types

In this section, we assess the descriptive characteristics of the different types of respondents we identify. Figure A6 shows the same kinds of analyses utilized in Figure 6 in the text but for various demographic and social characteristics of interest.

Figure A6: Demographics across Types



The figure shows kernel regressions (bandwidth = .1) of demographic characteristics across estimated ideologies for Downsians (black), Conversians (dark gray), and inattentive (light gray) respondents in the 2016 CCES.

Generally speaking, liberals and conservatives are more likely to be white, male, older, college educated and high income than Downsian moderates, Conversians, and inattentive respondents are. Also as expected, conservatives are more likely to attend church, while liberals are especially more likely to be young or have a college degree.

Focusing on non-ideologues with moderate estimated ideologies, Downsian moderates are more likely to be white, more likely to be high income, and less likely to attend church. Inattentive respondents are more likely than other groups to be Black or young.

Although demographics are correlated with our classifications, demographics would not necessarily allow one to accurately predict a respondent's type. For example, among our 2016 respondents, approximately 2.8 percent of those who are 30 years of age or older are classified as inattentive, while approximately 8.1 percent of those under 30 are classified as inattentive. So young people are much more likely to be inattentive, but only a small minority of young voters are inattentive. For these reasons, we would caution against researchers utilizing demographics as a proxy for whether survey respondents are Downsian, Conversian, or inattentive, as this would likely result in many misclassifications.

G Stability of Estimates

In this section we assess the stability of our estimates using data from the 2010-2014 Cooperative Congressional Election Panel Study (Schaffner and Ansolabehere, 2015). This data includes panel re-interviews for 9,500 respondents in the 2010, 2012, and 2014 waves of the Cooperative Congressional Election Studies. These respondents were asked the same questions as the respondents in our main results. We re-estimated our model for each of these three waves separately, and compared the estimates for each of these groups.

We are interested in the degree to which respondents retain the same "type" from wave to wave, particularly the degree to which respondents who are estimated to be Downsians in one wave are also classified as Downsians in other waves. Although we don't have a strong prediction for how often respondents should change types, we take stability as evidence of for the validity of the measurement. We are also interested in the degree to which respondent ideal points are stable. In particular, if our types are meaningful then the ideal points of Downsians should be more stable than the ideal points of non-Downsians.

Table A10 shows the percentage of respondents who are classified as Downsian or non-Downsian in 2010 and 2012. It shows that 82% of respondents are classified as Downsians in both years; and 94% of respondents classified as Downsian in 2010 are still classified as Downsian in 2012. Table A11 shows the same numbers for 2012 and 2014: 84% of respondents are classified as Downsian in both years, and among Downsians in 2012, 93% are still classified as Downsians in 2014. Our estimates across these years appear to be quite consistent when it comes to respondents classified as Downsians.

Table A10: Percent of respondents classified as Downsian in 2010 and 2012

	Downsian in 2012	Not Downsian in 2012
Downsian in 2010	82.2%	8.3%
Not Downsian in 2010	5.2%	4.3%

Table A11: Percent of respondents classified as Downsian in 2012 and 2014

	Downsian in 2014	Not Downsian in 2014
Downsian in 2012	84.4%	6.1%
Not Downsian in 2012	4.2%	5.3%

Figure A7 plots the estimated ideal points of non-Downsians and Downsians in 2012 and 2014. For non-Downsians, the correlation across these two time periods is 0.62. For Downsians, the correlation is 0.86. Doubtless some of this has to with the range of estimated ideal points, which is very compressed for non-Downsians. However the high degree of stability of the estimated ideal points of Downsians across two years is reassuring evidence that Downsians have meaningful policy views.

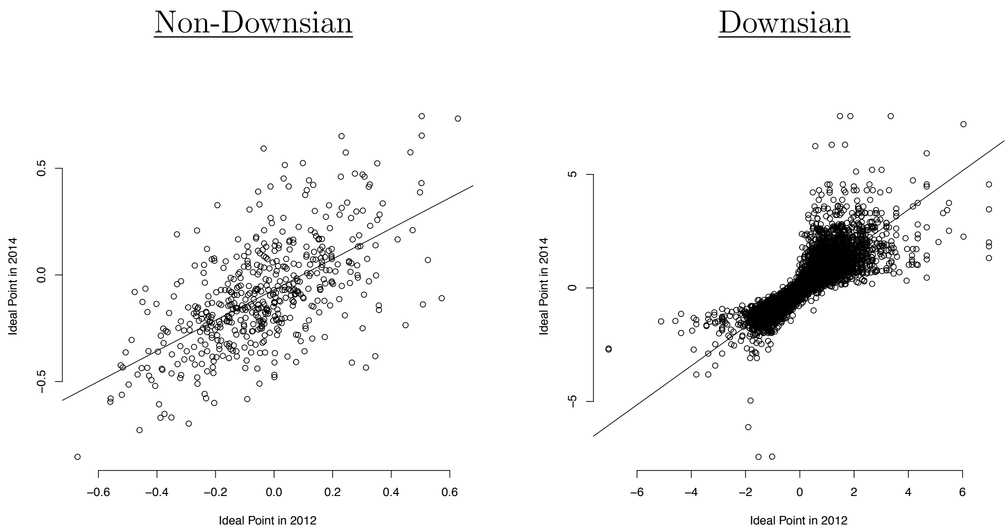


Figure A7: Stability of Estimated Ideal Points

Supplementary References

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

de Leeuw, Jan. 2011*a*. "A Horseshoe for Multidimensional Scaling." University of California Los Angeles, Department of Statistics Series.

de Leeuw, Jan. 2011*b*. "Quadratic and Cubic Majorization." University of California Los Angeles, Department of Statistics Series.

Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin. 1977. "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society: Series B (Methodological)* 39(1): 1–22.

Diaconis, Persi, Sharad Goel, and Susan Holmes. 2008. "Horseshoes in Multidimensional Scaling and Local Kernel Methods." *The Annals of Applied Statistics* 2(3): 777 – 807.

Hill, Mark O., and Hugh G. Gauch. 1980. "Detrended correspondence analysis: An improved ordination technique." *Vegetatio* 42(Oct): 47–58.

Imai, Kosuke, James Lo, and Jonathan Olmsted. 2016. "Fast Estimation of Ideal Points with Massive Data." *American Political Science Review* 110(4): 631–656.

Kendall, D. G. 1970. "A Mathematical Approach to Seriation." *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 269(1193): 125–134.

Schaffner, Brian, and Stephen Ansolabehere. 2015. "2010-2014 Cooperative Congressional Election Study Panel Survey." https://doi.org/10.7910/DVN/TOE8I1.

Shepard, Roger N. 1974. "Representation of Structure in Similarity Data: Problems and Prospects." *Psychometrika* 39: 373–421.