Online Appendix of
"Decomposing Political Knowledge: What is Confidence in Knowledge and Why it Matters"

Seonghui Lee
Aarhus University

Akitaka Matsuo
London School of Economics and Political Science

November 25, 2017

Contents

A Political Knowledge Questionnaire

1. Which group of foreign-born residents in the UK had the largest increase in population over the last decade?

1. Polish-born
2. Chinese-born
3. Indian-born
4. Nigerian-born

2. In what level is the current unemployment rate in the UK?

1. Lower than 7.0%
2. 7.0-8.5%
3. 8.6-10.0%
4. 10.1% or higher

3. Which party do you think this logo presents?

1. Conservative
2. Labour
3. Liberal Democrat
4. Scottish National Party
5. Plaid Cymru
6. Green Party

4. Which one of the following best describes the Alternative Voting (a.k.a. preferential voting) system?

1. Each voter has the chance to rank the candidates in order of preference.
2. Each voter votes for parties instead of for individual candidates.
3. Each voter has as many votes as there are choices, and can distribute those votes as desired.

5. Please choose the party or parties from the list below that are currently in the Cabinet. You may mark one or more parties.

1. Conservative
2. Labour
3. Liberal Democrat
4. Scottish National Party
5. Plaid Cymru
6. Green Party

6. From which of the following parties does the current Prime Minister come?

1. Conservative
2. Labour
3. Liberal Democrat
4. Scottish National Party
5. Plaid Cymru
6. Green Party

7. Identify the political party affiliation of Nick Clegg?

   1. Conservative
   2. Labour
   3. Liberal Democrat
   4. Scottish National Party
   5. Plaid Cymru
   6. Green Party

8. Identify the political party affiliation of Natalie Bennett?

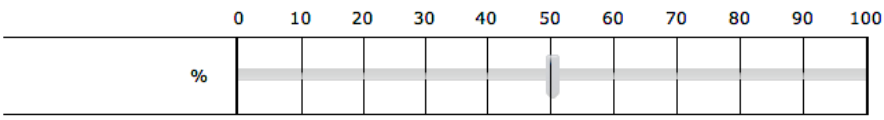
   1. Conservative
   2. Labour
   3. Liberal Democrat
   4. Scottish National Party
   5. Plaid Cymru
   6. Green Party

9. There are five parties below. Please rank-order the parties from the one that currently has the most seats to the least seats in the House of Commons.

   1. Conservative
   2. Labour
   3. Liberal Democrat
   4. Democratic Unionist Party
   5. Scottish National Party

** Each question is followed by the confidence rating question below. A pair of factual and confidence rating questions are shown on the same page.*

How confident are you that your answer is correct?



B Extended Model Descriptions

This section describes the technical details of the latent scale model we propose in the article. The theoretical framework of the paper necessitates independent measures of accuracy and confidence that correspond to the two conceptual constructs of political knowledge. A straightforward way to measure these two concepts might be to directly use survey responses , such as the sum of correct answers for accuracy and the average of confidence ratings for confidence. The problem of using these summary measures is that we have to assume that all questions are equally important to measure the retrieval accuracy or confidence in knowledge. This is obviously not always the case. For example, some inherently easy questions do not say much about the respondents' political knowledge while harder questions are likely to better distinguish respondents at various levels of informedness. We have shown in Figures 2 and 3, some political knowledge questions in our survey better differentiate the respondents than other items. This necessitates a methodology that detects characteristics of the items and assigns more weights on informative questions than less informative ones. The latent scaling models we employ here accomplish this.

We developed a model of latent scaling as an extension of Item Response Theory (IRT) model, which is widely used to estimate latent ability or traits (e.g. Delli Carpini and Keeter, 1993; Levendusky and Jackman, 2003). Our model estimates two latent scales of political knowledge – latent accuracy and latent confidence. Using a standard IRT model in which outcome variable is the correctness of the answer to each question, we estimate the latent accuracy for each respondent as a single measure. Similarly, using the self-reported confidence rating for each question as the outcome variable, we estimate the latent confidence scale through a method similar to the latent accuracy model. We include both latent scales in the specification of the latent confidence equation. This allows us to obtain a measure for the latent trait of confidence that is not directly caused by the latent accuracy and thus captures a single latent trait independent from the latent accuracy. The model includes two other equations in which latent scales are explanatory variables for political engagement and correct voting to test two hypotheses (H3 and H4). We simultaneously estimate latent scales and additional equations for hypotheses testing in order to deal with the issue of uncertainty inherent to the latent scale modeling.

B.1 Latent Accuracy Equation

The outcome variable in the latent accuracy equation is the correctness of the answer , a binary variable to indicate whether or not a respondent's answer is correct. A standard IRT model is appropriate to estimate this type of latent traits. In our application, we use a two-parameter IRT model (Jackman, 2009, 454-458). The outcome variable, Y_{ij}, is a binary variable for the correctness of respondent i's answer to j-th question. This Y_{ij} takes 1 if the answer is correct, and 0 otherwise. There is an auxiliary random variable, y_{ij}, which indicates the value of the binary outcome variable:

$$Y_{ij} = \begin{cases} 1 & \text{if } y_{ij} > 0 \\ 0 & \text{otherwise} \end{cases}$$

and, the value of y_{ij} is determined by the following formula:

$$\begin{aligned} y_{ij} &= \gamma_j \theta_i - \lambda_j + e_{ij} \\ e_{ij} &\sim N(0,1) \end{aligned} \tag{1}$$

where y_{ij} is a linear combination of a parameter at individual respondent level (θ_i), two parameters at item level (γ_j and λ_j), and a random error with the standard normal distribution (e_{ij}). The parameter of our main interest is θ_i which is an unobserved, latent measure of respondent i's retrieval accuracy. The larger value of θ_i indicates the higher retrieval accuracy. The intercept, λ_j, is an *item difficulty* of item j. Holding other parts of Equation 1 constant, it is less likely that y_{ij} takes positive value when the λ_j is larger. This implies that the item j is more difficult. The slope, γ_j, is an *item discrimination* parameter. When the value of γ_j is large, a small change in the accuracy level, θ_i, has a large effect on the value of y_{ij}. This implies that the probability of correctly answering a question drastically changes at a particular level of accuracy denoted by θ^*: most respondents with $\theta_i > \theta^*$ have a high chance to give the correct answer; most respondents with $\theta_i < \theta^*$ have a little chance to give the correct answer. As a result, an item with a larger γ_j better discriminates respondents on the left of the cut point against respondents on the right, in the sense that most of respondent with θ_i higher (or lower) than the cut point answer question correctly (or incorrectly).[1]

B.2 Latent Confidence Equation

In our survey, each factual knowledge question is followed by a question about the respondent's confidence in the correctness of the answer. The outcome variable in the confidence equation is the reported confidence rating, C_{ij}. We assume that there is a general, individual-level latent trait of confidence in political knowledge for a respondent i, and this latent confidence, denoted δ_i, has a direct impact on the reported confidence rating for her/his answer to each question. The confidence equation

[1] This point might be easily understood by presenting the model in another form. The model can be presented as a probit model with a latent, unobserved variable θ_i:

$$\begin{aligned} \Pr(Y_{ij} = 1|\theta_i, \gamma_j, \lambda_j) &= \Pr(y_{ij} > 0|\theta_i, \gamma_j, \lambda_j) \\ &= \Phi(\gamma_j \theta_i - \lambda_j) \end{aligned}$$

where Φ is the standard normal CDF.

has a specification similar to the latent accuracy model (Equation 1), but C_{ij} is directly measurable from survey questions on a continuous scale.[2]

The confidence equation is specified as a linear model with a random error term:

$$C_{ij} = \beta_{1j}\delta_i + \beta_{2j}\theta_i - \alpha_j + \epsilon_{ij} \tag{2}$$
$$\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$$

where two latent measures of political knowledge are simultaneously included. The latent accuracy, θ_i, is included in this equation to capture its effect on item-level confidence rating, C_{ij}. It is worth highlighting that by including both terms, the estimated confidence level δ_i becomes the latent confidence level *not explained* by the retrieval accuracy. This separation of δ_i from θ_i assures the independence of latent confidence from latent accuracy, allowing us to explore the confidence in knowledge as a latent trait after accounting for its contribution to retrieval accuracy.

The resemblance of Equations 2 and 1 allows simple interpretation of item-level parameters in Equation 2. Each of the two latent scales at the individual level is multiplied by an item-specific slope coefficient (β_{1j} or β_{2j}). The meanings of these slopes are similar to the discrimination parameter γ_j in the accuracy equation: The larger β_{1j} (or β_{2j}) is, the larger the effect of the latent confidence (or accuracy) on the reported confidence rating for item j. When β_{1j} is larger than β_{2j}, the impact of latent confidence on the reported confidence for item j is larger than the impact of the latent accuracy.[3] The intercept, α_j, is analogues to λ_j (the item difficulty parameter in the latent accuracy equation): when an item j has a larger α_j, it indicates that the item is one that is "feeling difficult" so that respondents tend to report lower confidence ratings on average.

B.3 Other Outcome Variables Explained by Latent Scales

Our hypotheses regarding consequences of political knowledge (H3 and H4) contend that retrieval accuracy and confidence-in-knowledge will explain other outcome variables, such as political engagement and informed vote choice. To test these hypotheses, we use the latent accuracy and confidence scales as predictors of these variables.

In the equations where the latent traits are explanatory variables, we include both latent scales in the same equation and compare the sizes of effects of these two scales on the outcome variable (political engagement or informed vote choice). The dependent variable in the political engagement model is *political discussion*, measured with a four-point scale from never discuss to frequently discuss. For the informed vote choice model we employ the concept of "correct voting" proposed by Lau and Redlawsk

[2] In the actual estimation, C_{ij} is a logit transformation of the survey answer ranging from 0 to 100.

[3] We claim the direct comparability between β_{1j} and β_{2j} based on the identical distributions of θ_i and δ_i. See Section B.4.1.

(1997). Among various ways to operationalize the concept, we utilize the one related to the "prospective policy-based considerations" used in Lau et al (2014) to measure whether a respondent makes a decision by voting for a party whose position is the most approximate to the person's position in the left-right ideological spectrum. The specification of the political discussion model is a normal linear model and the informed choice is a probit model.

To test them, we include both the latent accuracy and confidence scales as the predictors of the outcome variables. Since the latent scales have no single representative value as they are measured as a distribution, it might be an intuitive solution to use summary measures – such as mean of median of distributions – to estimate these models. However, this solution ignores the uncertainty around the latent scales and the results can be biased (Hollyer, Rosendorff and Vreeland, 2014; Junker, Schofield and Taylor, 2012). Therefore, we instead incorporate our two latent trait models into a unified model and jointly estimate this system of models exploiting the flexibility of Bayesian modeling (c.f. Fox and Glas, 2003).

B.4 Model Identification for Latent Scales

Latent scale models have identification problem because models are identified up to a linear transformation even in the one dimensional scaling case (c.f. Fox, 2010, Chapter 4), and this issue could be more complicated under the multidimensional latent scaling because of the issue of rotation. Rivers (2003) discusses the case of multidimensional latent models for the IRT modeling, with particular focus on the ideological point estimation using the roll-call votes as the input of the model.

Despite some similarities between our identification problem and that of higher dimensional ideal point estimation, they are essentially different. The difficulty of multidimensional ideal point estimation for roll-call data is that there is only one outcome variable. Therefore, to identify the model, both the dispersion of ideal points in each dimension and the rotation of the scale have to be properly addressed. In contrast, our model have two outcome variables. Because there are more than one outcome variables, we do not need to worry about the issue of rotation, and the model will be identified as long as the dispersion of each of latent scales are well-defined. This can be achieved by normalization (global identification) or by fixing the variance (local identification). In the following, we present the strategy for both global and local identification, followed by the comparison of estimates.

B.4.1 Local Identification

The local identification can be achieved with a relatively simple strategy that assigns the dispersion of two latent scales, θ and δ, separately. This is parallel to the strategy to identify the unidimensional ideal point estimation using the IRT model. Following Jackman (2001), we identify the model by imposing the standard normal distribution prior on the latent scales ($\theta \propto N(0,1)$ and $\delta \propto N(0,1)$).[4] We set diffuse prior distributions for other parameters.

The posterior density is obtained through the Markov Chain Monte Carlo simulation implemented in `jags 4.1.0`. We run 20,000 iteration for two chains with thinning of 10 after 5,000 burn-in. The MCMC diagnostics show no sign of non-convergence. The `JAGS` code of this joint model is available in Online Appendix B and summary and visualized convergence diagnostics is presented in Online Appendix G.

Because of this prior specification, the estimated two latent scales, θ and δ, have similar posterior distributions. When we calculate the median value of each individual's posterior θ (or δ) samples, the average across all individuals is virtually zero and the standard deviation is one. We believe that this close resemblance between empirical distributions of these two parameters provides the foundation for testing and comparing relative importance of the two latent scales against each other.

B.4.2 Global Identification

Another possible strategy is to globally identify the model by settingset restrictions on parameters. For instance, in the case of unidimensional IRT models for scaling legislators' ideological points using roll-call records as input data, the model is typically identified by setting the location of two ideologically extreme legislators, one on the left and the other on the right. Rivers (2003) refer to this type of restriction as the "Kennedy-Helms" restriction, based on the names of two Senators used by Poole and Rosenthal (2000). By setting such restriction, the model is identified because the positions of other members of the legislature can be measured as the relative distance from these two extreme points. By this specification, outcome measures are "normalized" because most other members are located between these two extreme legislators. For the two dimensional IRT case, a model can be identified by setting the restrictions on three ideal points (Jackman, 2001).

Although our model also has two latent scales, the identification problem is substantially different from the case of two dimensional IRT model described above. For our latent accuracy model, the identification can be achieved by setting proper restrictions on the accuracy equation (Equation 1). As is the case for the unidimensional IRT, we identify the model by setting two fixed points on the dimension. The identification of the second latent scale, which we plan to conduct through the parameter restrictions on the accuracy equation (Equation 2), is achieved by setting two points on the confidence dimension. For each of latent scales, we set parameter restriction only on that scale and not on the other (e.g. for the latent accuracy, we *did not* set the restrictions on latent confidence parameter values for

[4] Rivers (2003) is critical about the strategy of using the dispersion parameter of priors to identify the model, because with that prior specification it is not clear "what information is actually being brought to bear on the problem (p. 7)." Although this may be a valid criticism, from the practical standpoint, this specification provides posterior convergence as well as all reasonable posterior distributions without normalizing the posterior.

those who we set the fixed value for their latent accuracy parameter), because we do not have to set fixed values for the other scale for identification purpose.[5]

The question that still remains is the selection of survey respondents to be used for this parameter restriction. In the case of "Kennedy-Helms" restriction, two Senators are chosen by researchers' substantive knowledge about Senators who are ideologically extreme in the chamber. However, in our survey data , we do not have such a-priori knowledge about survey respondents. Therefore, we chose the extreme points based on actual survey responses. In our latent accuracy model, we randomly chose one respondent who gave no correct answer and another respondent who gave correct answers for all questions. We set the value of the least accurate respondent's latent accuracy to -1 and that of the most accurate to 1. In our latent confidence model, we chose two provided with the highest and lowest average confidence rating. Again, we set the value of the least confident respondent's latent confidence to -1 and that of the most confident to 1. Since we do not have much prior knowledge on the variances of the latent scales, we use diffuse priors. With this specification, we conducted the MCMC sampling of 20,000 iteration for two chains with thinning of 10 after 5,000 burn-in. The post sample diagnostics indicates the model convergence. As we intended, the correlation between two scales are very low ($r^2 = .06$).

Below we presents results from the globally identified models. Tables B.1 are B.2 correspond to Tables 2 and 3 in the main text. The sizes of coefficients for latent scales are different than those from the locally identified models because of the smaller variance of latent scales. However, the direction of the coefficients and the frequentist equivalent of the significance (or whether the 95% Bayesian HPD overlaps with 0) does not change between the locally and globally identified models. Consequently, our substantive interpretation of the results remains the same. In the next subsection, we show that the results from these two identification strategies are essentially equivalent after rescaling of the latent scales from the globally identified models.

B.4.3 Comparison of the Results from Two Identification Strategies

Our local identification strategy, which follows Jackman (2001), identifies the model by imposing an informed prior on the distributions of two latent scales. We use the standard normal distribution for both scales, and because of that, the dispersions of two latent scales are almost identical, with mean zero and variance one. We consider that this enables direct comparison between the two latent scales' impact on other variables (in models where the latent scales are input) and the effect of other variables on the two latent scales (in models where the latent scales are output).

This is not the case for the globally identified models. Although we may try "normalizing (Rivers, 2003, p.7)" the scales with an intention to bound the distributions of each scale in $(-1, 1)$, these restrictions do not necessarily return the two scales distributed in the similar way In our globally

[5] Also, we do not have much prior information about the position of their location in the other dimension.

Table B.1: Consequences of Political Knowledge: Political Engagement (Globally identified)

	Dependent variable: Political Discussion		
	Model 1	Model 2	Model 3
Nubmer of Correct Answer	0.070 (0.012, 0.127)		
Latent Accuracy		0.151 (0.011, 0.313)	0.166 (0.006, 0.343)
Latent Confidence		0.597 (0.395, 0.830)	0.592 (0.393, 0.840)
Accuracy x Confidence			0.058 (-0.472, 0.584)
Gender (Male)	0.046 (-0.067, 0.158)	-0.016 (-0.129, 0.096)	-0.017 (-0.127, 0.095)
Education	0.059 (0.012, 0.106)	0.054 (0.009, 0.100)	0.054 (0.008, 0.102)
Party Identifier	0.504 (0.328, 0.679)	0.400 (0.223, 0.572)	0.401 (0.222, 0.581)
Age	-0.753 (-3.149, 1.644)	-1.001 (-3.094, 1.513)	-1.016 (-3.497, 1.118)
Age^2	0.159 (-2.453, 2.772)	0.614 (-2.126, 2.926)	0.673 (-1.747, 3.339)
Intercept	2.486 (1.904, 3.068)	2.635 (2.016, 3.153)	2.648 (2.133, 3.246)
Observations	790	790	790
R^2	0.071		

Model 1 correspond to the models in Table 3 of the text.
Models 2 and 3 are the results from the globally identified model.

identified models , the variances of the two distributions are substantially different: The variance of the median of posterior samples for the accuracy scale is 0.443 and that for the confidence scale is 0.346. To make the variance of two scales similar to each other and adjust the estimated coefficients accordingly, we calculate the variance-covariance matrix of the latent scales in each iteration, and modify the coefficients using the inverse variance-covariance matrix. From that, we regenerate the results reported in Tables A.1 and A.2 to compare them with those from locally identified models in Tables 2 and 3 in main text. Tables B.3 and B.4 below report the results for the main variables of interests only (i.e., the effects of the two latent scales on political discussion and informed vote choice.) The tables indicate that the results from the globally identified models are almost identical to those from the locally identified models.

Table B.2: Consequences of Political Knowledge: Voting Correctly (Globally identified)

	Dependent variable: Correct Voting		
	Model 1	Model 2	Model 3
Nubmer of Correct Answer	0.187		
	(0.085, 0.289)		
Latent Accuracy		0.418	0.514
		(0.153, 0.749)	(0.218, 0.908)
Latent Confidence		-0.249	-0.407
		(-0.604, 0.073)	(-0.816, -0.059)
Accuracy x Confidence			1.501
			(0.529, 2.884)
Gender (Male)	0.007	0.026	0.018
	(-0.188, 0.201)	(-0.183, 0.223)	(-0.189, 0.222)
Education	0.090	0.096	0.098
	(0.007, 0.174)	(0.017, 0.184)	(0.015, 0.183)
Party Identifier	0.710	0.751	0.824
	(0.349, 1.072)	(0.398, 1.140)	(0.461, 1.222)
Age	3.082	4.121	2.781
	(-1.165, 7.329)	(0.332, 8.686)	(-0.855, 7.139)
Age^2	-2.251	-3.486	-2.049
	(-6.827, 2.326)	(-8.428, 0.570)	(-6.697, 2.017)
Intercept	-2.332	-2.613	-2.402
	(-3.402, -1.262)	(-3.777, -1.630)	(-3.506, -1.421)
Observations	790	790	790
AIC	965.370		

Model 1 correspond to the models in Table 3 of the text.
Models 2 and 3 are the results from the globally identified model.

Table B.3: Comparison of the Two Identification Strategies: Political Engagement Model

	Dependent variable: Political Discussion			
Identification	Global		Local	
	(1)	(2)	(3)	(4)
Latent Accuracy	0.071	0.073	0.077	0.076
	(0.005, 0.138)	(0.003, 0.140)	(0.009, 0.145)	(0.010, 0.145)
Latent Confidence	0.206	0.203	0.204	0.202
	(0.147, 0.263)	(0.141, 0.266)	(0.142, 0.266)	(0.139, 0.265)
Accuracy x Confidence		0.009		0.006
		(-0.068, 0.082)		(-0.072, 0.080)
Interaction Term	No	Yes	No	Yes
Observations	790	790	790	790

Models 1 and 2 are the results from Globally Identified Models. Coefficients are adjusted through the standardization of θ and δ.
Models 3 and 4 are the results from Locally Identified Models.

Table B.4: Comparison of the Two Identification Strategies: Correct Voting Model

	Dependent variable: Voting for a party that is ideologically closest to the respondent			
Identification	Global		Local	
	(1)	(2)	(3)	(4)
Latent Accuracy	0.193	0.225	0.196	0.221
	(0.072, 0.315)	(0.097, 0.360)	(0.076, 0.320)	(0.091, 0.357)
Latent Confidence	-0.087	-0.139	-0.087	-0.141
	(-0.201, 0.025)	(-0.266, -0.022)	(-0.198, 0.020)	(-0.268, -0.016)
Accuracy x Confidence		0.222		0.223
		(0.083, 0.383)		(0.075, 0.382)
Interaction Term	No	Yes	No	Yes
Observations	790	790	790	790

Models 1 and 2 are the results from Globally Identified Models. Coefficients are adjusted through the standardization of θ and δ.
Models 3 and 4 are the results from Locally Identified Models.

B.5 Political Knowledge as Outcome Variable

In our two models where political knowledge is used as outcome variable the two latent scales are the linear combination of variables that affect political knowledge, with normally-distributed random errors. We estimate two linear models: one for latent accuracy and another for latent confidence as the outcome variable:

$$\begin{aligned} \theta_i|\mathbf{X}_i,, \xi_\theta, \mu_\theta &\sim N(\mathbf{X}_i\boldsymbol{\xi}_\theta, \mu_\theta) \\ \delta_i|\mathbf{X}_i,, \xi_\delta, \mu_\delta &\sim N(\mathbf{X}_i\boldsymbol{\xi}_\delta, \mu_\delta) \end{aligned} \tag{3}$$

where $\mathbf{X}$ is a vector of explanatory variables that are common in both accuracy and confidence equations. $\boldsymbol{\xi}_\theta$ ($\boldsymbol{\xi}_\delta$) is a vector of regression coefficients for θ (δ), and μ_θ (μ_δ) is the variance of normally distributed errors. The explanatory variables include Political Interest, Education, Party Identification, Male, and Age. To incorporate the uncertainty around the latent scale estimates, we employ a strategy advocated by Armstrong II et al. (2014, 282-6), in which a model is estimated repeatedly for each value of 2,000 MCMC samples, and the distribution of the point estimates for each parameter is used to calculate a confidence interval. For comparison purpose, we also estimate the models where we use the survey measures directly from the survey responses as the outcome variables, such as the number of correct answers (for accuracy) and the average confidence ratings (for confidence). These two variables of the observed measures are standardized. We present the results in Online Appendix E.

B.6 Jags Code for the Model

The following is the JAGS code for our main model.

```
model {
  for(i in 1:N){
    for(j in 1:M){
      #latent accuracy equation update
      Y[i, j] ~ dbern(prob[i,j])
      probit(prob[i,j]) <- (theta[i]*gamma[j] - lambda[j])

      #latent confidence equation update
      conf.data[i,j] ~ dnorm(conf.data.hat[i,j],tau.e)
      conf.data.hat[i,j] <- (beta1[j]*delta[i] + beta2[j]*theta[i] - alpha[j])

    }
    ## latent scales as input model
    # update a model of political discussion
    poldisc[i]  ~ dnorm(theta[i]*d1[1]+delta[i]*d1[2]+input.data[i,]%*%d1[3:8],tau.e1)
    # update a model of correct voting
    vote.correct[i] ~ dbern(prob.vote.correct[
    probit(prob.vote.correct[i]) <- (theta[i]*d2[1]+delta[i]*d2[2]+input.data[i,]%*%d2[3:8])
  }

  # update the individuals capability papameter
  for(i in 1:N){
    theta[i] ~ dnorm(0, 1)
    delta[i] ~ dnorm(0, 1)
  }

  # d1 and d2: coefficient for polidisc model
  for (i in 1:8){
    d1[i] ~ dnorm(mu.d1[i],tau.d1[i])
    d2[i] ~ dnorm(mu.d2[i],tau.d2[i])
  }

  ## draw item parameters for confidence equation
  for(j in 1:M){
    beta1[j] ~ dnorm( mu.beta1, tau.beta1 )
    beta2[j] ~ dnorm( mu.beta2, tau.beta2 )
    alpha[j] ~ dnorm( mu.alpha, tau.alpha )
  }

  ## draw item parameters for accuracy equation
  for(j in 1:M){
    lambda[j] ~ dnorm( mu.lambda, tau.lambda )
    gamma[j] ~ dnorm( mu.gamma, tau.gamma)
  }

  ## prior setup
  mu.beta1 ~ dnorm(1, 0.0001)
  tau.beta1 <- pow(sigma.beta1, -2)
  sigma.beta1 ~ dunif(0,100)
```

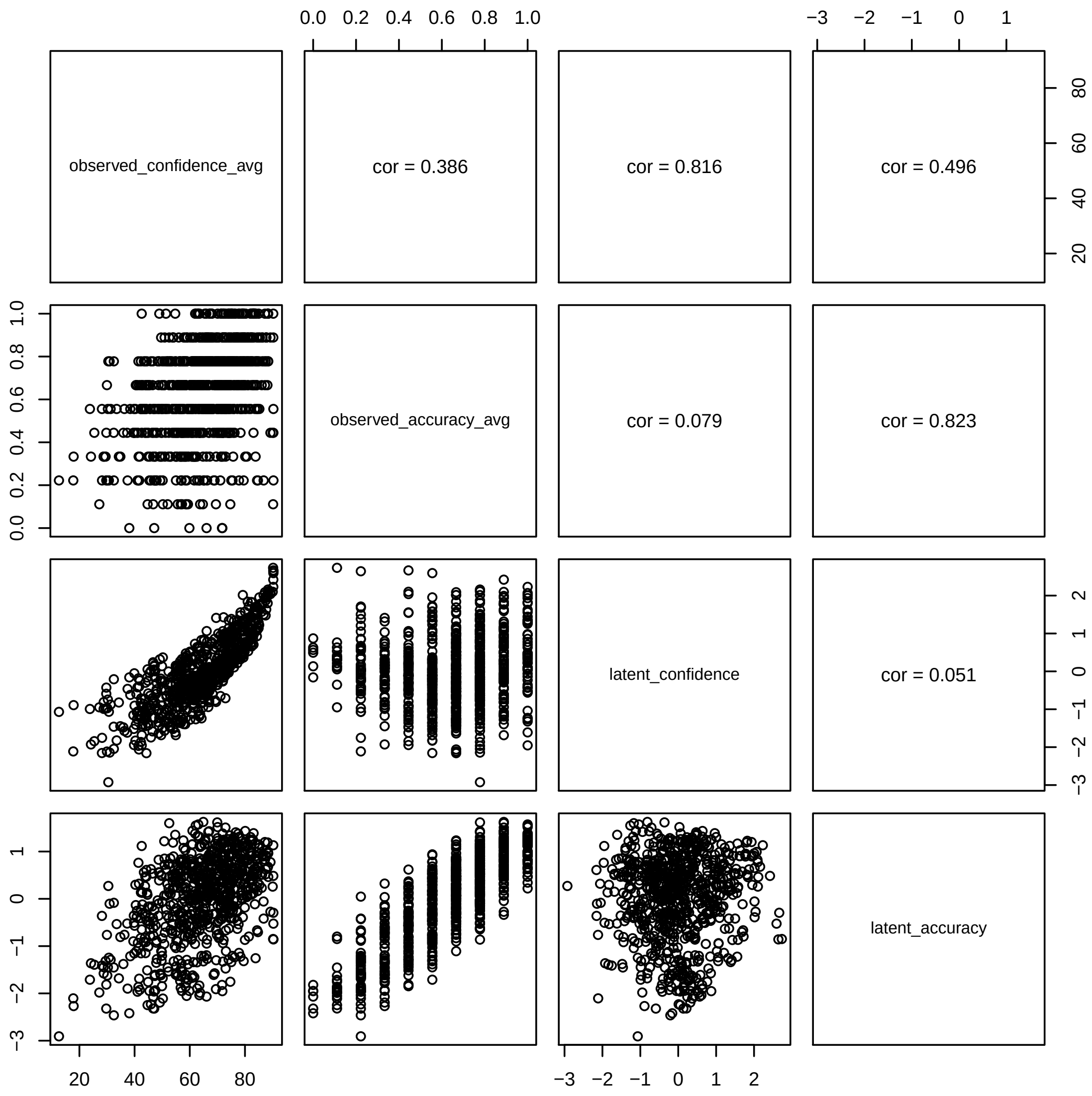
```
  mu.beta2 ~ dnorm(1, 0.0001)
  tau.beta2 <- pow(sigma.beta2, -2)
  sigma.beta2 ~ dunif(0,100)

  mu.alpha ~ dnorm(0, 0.0001)
  tau.alpha <- pow(sigma.alpha, -2)
  sigma.alpha ~ dunif(0,100)

  tau.e <- pow(sigma.e,-2)
  sigma.e ~ dunif(0,100)

  mu.lambda ~ dnorm(0, 0.0001)
  sigma.lambda ~ dunif(0,100)
  tau.lambda <- pow(sigma.lambda, -2)
  mu.gamma ~ dnorm(0, 0.0001)
  sigma.gamma ~ dunif(0,100)
  tau.gamma <- pow(sigma.gamma, -2)

  for(i in 1:8){
    mu.d1[i] ~ dnorm(0, 0.0001)
    sigma.d1[i] ~ dunif(0,100)
    tau.d1[i] <- pow(sigma.d1[i], -2)
    mu.d2[i] ~ dnorm(0, 0.0001)
    sigma.d2[i] ~ dunif(0,100)
    tau.d2[i] <- pow(sigma.d2[i], -2)
  }

  tau.e1 <- pow(sigma.e1,-2)
  sigma.e1 ~ dunif(0,100)

}
```

C Correlations Plots between Observed and Latent Scales for Accuracy and Confidence in Political Knowledge

Figure C.1 is a correlation matrix plot of four scales of political knowledge: two measures for accuracy and confidence based on survey responses (the number of correct answers and the mean confidence ratings for individual respondents) and the latent scales for accuracy and confidence in political knowledge obtained from our latent scaling models. The panels above the diagonal show the correlations between two of the scales and the panels below the diagonal illustrate the scatter plots.

The latent and observed scales of accuracy (or confidence) are in close resemblance. Both of the correlation coefficients between the latent and observed scales for accuracy and confidence are higher than .8 and the scatter plots illustrate almost linear relations between them. These results might lead to an impression that the observed scales can be used as the substitute for the latent one. However, this is only partly true (see Online Appendix E for more discussion). There exists a moderate-to-weak correlation between *observed* accuracy and *observed* confidence ($r^2 = .386$). This is probably because the reported confidence ratings are not only influenced by individual respondents' latent accessibility of information (latent confidence), but also by the latent ability of retrieval accuracy. We have taken this point into account in our latent confidence model, and as the result, the correlation between the *latent* accuracy and *latent* confidence scales is only 0.051. We aimed to estimate the latent confidence δ_i that is not directly explained by the retrieval accuracy, so that the resulting latent confidence scale captures a single latent trait for the accessibility and availability of relevant information independently from the latent accuracy. For that goal, our latent scale of confidence serves well.

Figure C.1: Correlation Matrix Plot: Observed and Latent Scales of Accuracy and Confidence in Political Knowledge



D Descriptive Statistics

Table D.1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Age	44.294	14.805	15	79	790
Male	0.576	0.495	0	1	792
Education	4.095	1.167	1	5	792
Party Identifier	0.890	0.313	0	1	792
Political interest	3.237	0.835	1	4	792
Political discussion	2.902	0.792	1	4	792
Politics complicated	2.946	0.906	1	5	792
Number of correct answers	5.693	2.001	0	9	792
Average confidence rating	63.989	23.254	0	100	792

Figure D.1: Correlation among variables at the individual level

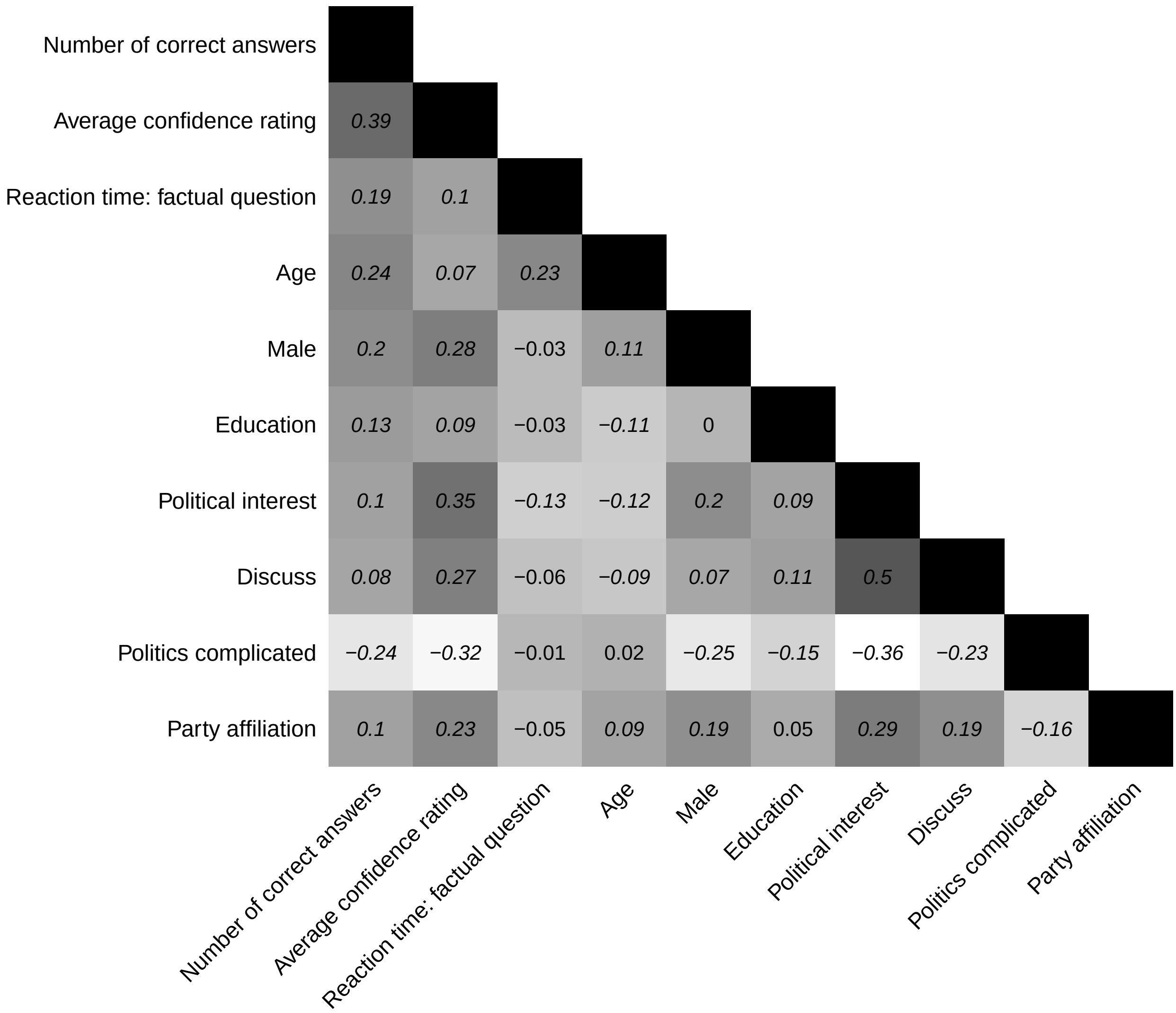
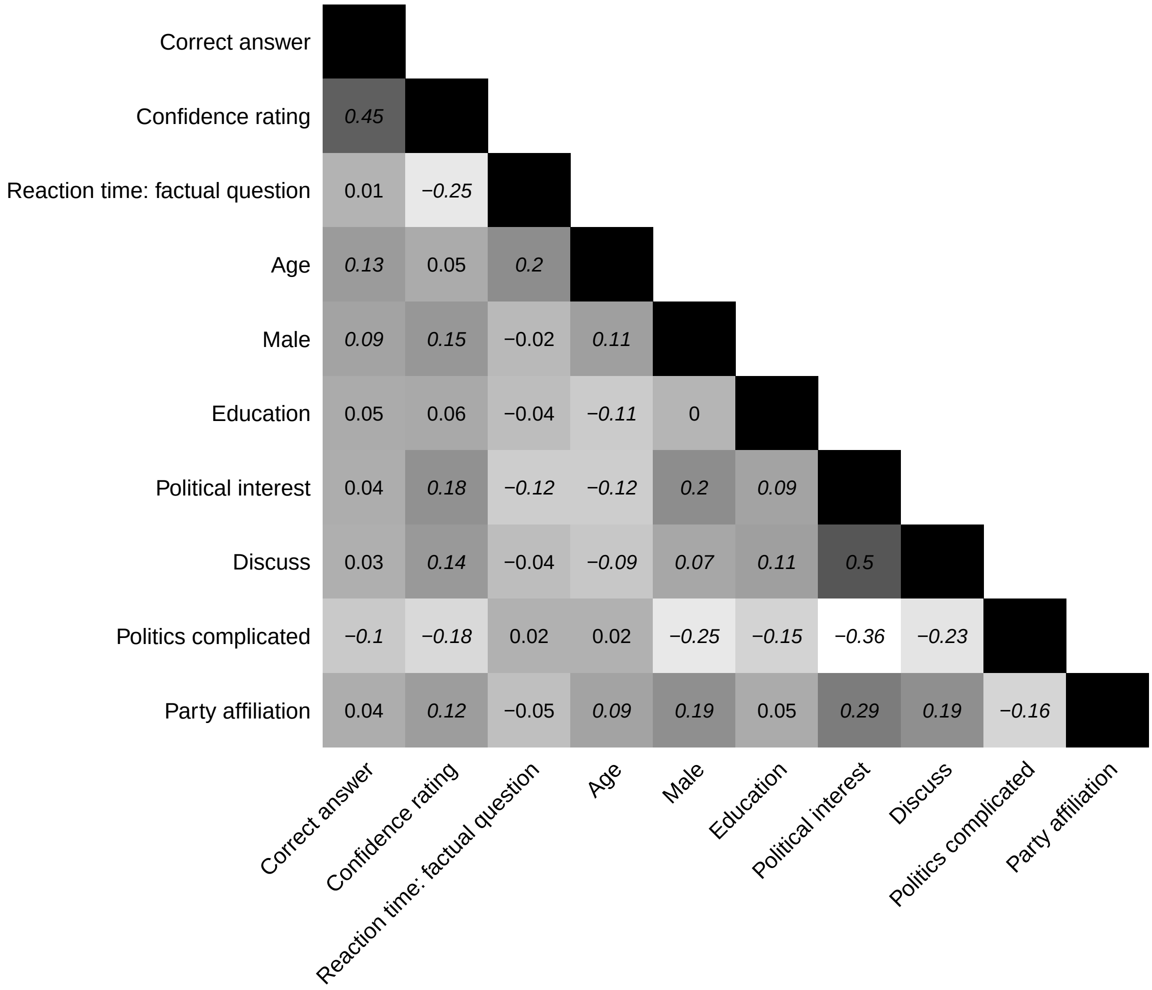


Figure D.2: Correlation among variables at the item level



E Additional Model Estimates

E.1 Political Knowledge as Explanatory Variable: Models Including Average Confidence Rating

In this subsection, we reproduce Tables 2 and 3 in the text with an additional model (Model 1b). The added model include standardized average of confidence ratings across items (Average Confidence) as well as the Number of Correct Answer. Table E.1 below is the replication of the political engagement model and Table E.2 is the replication of the correct voting model. Comparing the Model 1b to Model 3 in political engagement models (Table E.1), the size of the effect of Average Confidence is similar to that of Latent Confidence. This is consistent with our finding that: "Individuals with a higher level of confidence-in-knowledge are more likely to be active in political discussions than individuals with the same level of retrieval accuracy." However, there is an interesting difference in the effects of retrieval accuracy between Model 1b and Model 3: The effect of Number of Correct Answers is not significant in Model 1b while the effect of Latent Accuracy is significant in Model 2. This happens because there is a moderate correlation between the two observed measures ($r^2 = 0.39$), and when both are included in the same model, the effect of retrieval accuracy is masked. In the article we proposed a joint model of two latent scales to explore the two-dimensional conceptual space of accuracy and confidence. We believe that this finding would be another justification of using the latent measures of knowledge.

We do not see such difference between the results using observed measures and those using latent scales in the Correct Voting models (Table E.2). The estimates of the observed and latent accuracy are roughly the same. This implies that the associations between accuracy and correct voting is robust. The higher accuracy in political knowledge an individual has, the more likely she makes a correct decision in choosing a party to vote for; and this is not mitigated by the confidence in knowledge.

Table E.1: Consequences of Political Knowledge: Political Engagement

	Dependent variable: Political Discussion		
	Model 1	Model 1b	Model 3
Nubmer of Correct Answer	0.070 (0.012, 0.127)	-0.002 (-0.061, 0.058)	
Average Confidence		0.210 (0.150, 0.269)	
Latent Accuracy			0.076 (0.010, 0.145)
Latent Confidence			0.202 (0.139, 0.265)
Accuracy x Confidence			0.006 (-0.072, 0.080)
Gender (Male)	0.046 (-0.067, 0.158)	-0.033 (-0.144, 0.078)	-0.015 (-0.130, 0.095)
Education	0.059 (0.012, 0.106)	0.054 (0.008, 0.100)	0.055 (0.009, 0.100)
Party Identifier	0.504 (0.328, 0.679)	0.369 (0.195, 0.544)	0.403 (0.225, 0.573)
Age	-0.753 (-3.149, 1.644)	-1.138 (-3.468, 1.193)	-0.979 (-3.052, 1.363)
Age^2	0.159 (-2.453, 2.772)	0.669 (-1.873, 3.211)	0.584 (-1.933, 2.839)
Intercept	2.486 (1.904, 3.068)	2.728 (2.158, 3.297)	2.635 (2.073, 3.164)
Observations	790	790	790
R^2	0.071	0.124	

Models 1 and 3 correspond to the models in Table 2 of the text. Model 1b is added.
Models 1 and 1b: Ordinary linear regression for comparison, Model 3: Results from the joint model.
Numbers in the parentheses indicate the upper and lower bounds of 95 percent confidence intervals (Models 1 and 1b) and credible intervals (Model 3).

Table E.2: Consequences of Political Knowledge: Voting Correctly

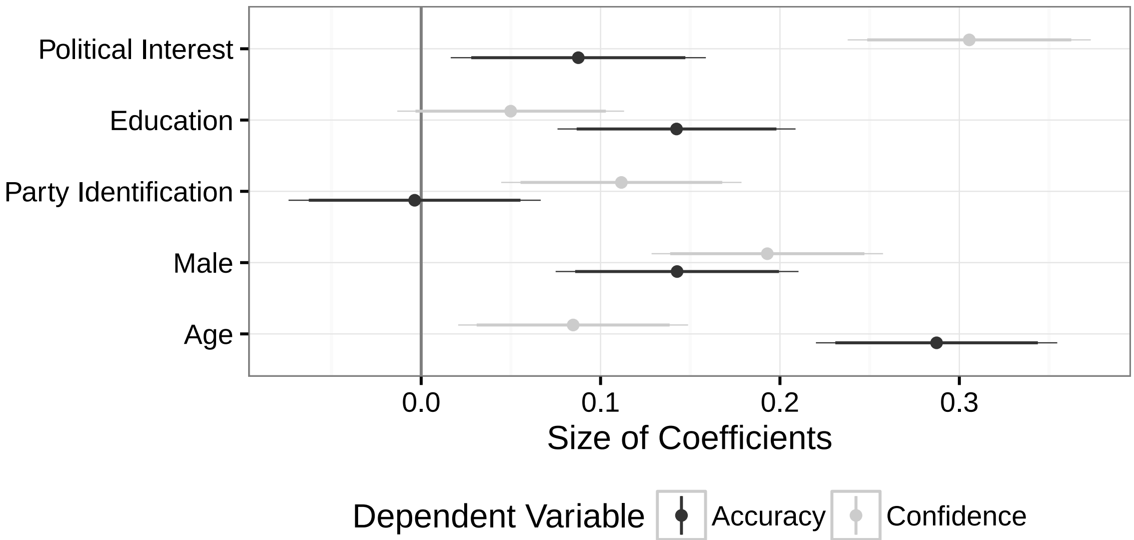
	Dependent Variable: Voting for a party that is ideologically closest to the respondent		
	Model 1	Model 1b	Model 2
Nubmer of Correct Answer	0.187	0.230	
	(0.085, 0.289)	(0.120, 0.340)	
Average Confidence		-0.116	
		(-0.225, -0.007)	
Latent Accuracy			0.221
			(0.091, 0.357)
Latent Confidence			-0.141
			(-0.268, -0.016)
Accuracy x Confidence			0.223
			(0.075, 0.382)
Gender (Male)	0.007	0.046	0.015
	(-0.188, 0.201)	(-0.153, 0.246)	(-0.187, 0.221)
Education	0.090	0.092	0.097
	(0.007, 0.174)	(0.009, 0.175)	(0.016, 0.182)
Party Identifier	0.710	0.775	0.805
	(0.349, 1.072)	(0.406, 1.144)	(0.436, 1.198)
Age	3.082	3.327	2.931
	(-1.165, 7.329)	(-0.928, 7.581)	(-0.941, 6.793)
Age^2	-2.251	-2.557	-2.183
	(-6.827, 2.326)	(-7.145, 2.030)	(-6.399, 1.914)
Intercept	-2.332	-2.463	-2.401
	(-3.402, -1.262)	(-3.542, -1.385)	(-3.474, -1.374)
N	790	790	790
AIC	965.370	962.968	

Models 1 and 2 correspond to the models in Table 3 of the text. Model 1b is added.
Models 1 and 1b: Probit model (MLE), Model 2: Results from the joint model. Numbers in the parentheses indicate the upper and lower bounds of 95 percent confidence intervals (Models 1 and 1b) and credible intervals (Model 2).

E.2 Political Knowledge as Outcome Variable: Models Using Observed Measures of Accuracy and Confidence

Figure E.1 presents the results from models in which the outcome variables are observed measures for accuracy (number of correct answers) and confidence (average confidence ratings). The point estimates are similar to those in Figure 8 in the text (except for the effect of Male), but have larger confidence intervals.

Figure E.1: The Effects of Political Interest and Education on Two Latent Traits



Note: The point indicates the mean of coefficients estimates. Solid lines indicate the confidence intervals, thick lines for 90 percent confidence intervals and thin lines for 95 percent confidence intervals. All variables are standardized.

F Comparison with British Election Study Data

The survey respondents in our study are opt-in sample from the SSI's online pool. In this section, we provide comparisons between our SSI sample and a probability based sample in a large-scale political survey in their key demographic and psychological variables. The representative survey we use as a point of comparison is the Wave 1 of 2014-2017 British Election Internet Panel Study (BES). Our study was fielded in November 2013 and the BES Wave 1 was fielded between 20 February and 9 March, 2014.

Table F.1 presents the comparisons of averages in five variables that are included in both surveys. Overall, the differences between our samples and BES samples are not much of concern. The largest difference we have found is the age of respondents: The SSI sample is apparently younger (44.3) than the BES's (51.6 or 54.5). However, note that the mean age of voting population in the UK is 47.7, falling in between the mean age of the SSI and the BES. The proportion of party identifier and level of education are close. The level of political interest is slightly higher for our samples than BES.

Table F.1: Comparison of SSI and BES Samples

Variable	Mean SSI	Mean BES	Weighted Mean BES	N (SSI)	N (BES)	Min	Max
Age	44.294	51.557	54.499	790	30235	18	95
Male	0.575	0.505	0.529	790	30235	0	1
Education	4.094	3.863	3.842	790	29377	1	5
Party Identifier	0.890	0.862	0.870	790	30157	0	1
Political Interest (rescaled)	0.746	0.690	0.695	790	30036	0	1

Note: We use ”Attention to politics” question in the BES (“How much attention do you generally pay to politics?”) to make comparison with the typical ”Political interest” question in the SSI. Since the “Attention to politics” and “Political interest” question used different response scales (BES with 11-point, SSI with 4-point scale), we rescale these questions to make the range of variable from 0 to 1. The Party Identification question is a dichotomous variable where the value takes 1 when respondents are identified with any of political parties in the UK or think of themselves as closer to any of the parties. For BES, we calculate both unweighted and weighted averages (using a weighting variable `w1core`).

References

Armstrong II, David A., Ryan Bakker, Royce Carroll, Christopher Hare and Howard Rosenthal. 2014. *Analyzing Spatial Models of Choice and Judgment with R*. Boca Raton, FL: CRC Press.

Delli Carpini, Michael and Scott Keeter. 1993. "Measuring Political Knowledge: Putting First Things First." *American Journal of Political Science* 37(4):1179–1206.

Fox, Jean-Paul. 2010. *Bayesian item response modeling: Theory and applications*. Springer.

Fox, Jean-Paul and Cees A. W. Glas. 2003. "Bayesian Modeling of Measurement Error in Predictor Variables Using Item Response Theory." *Psychometrika* 68(2):169–191.

Hollyer, J. R., B. P. Rosendorff and J. R. Vreeland. 2014. "Measuring Transparency." *Political Analysis* 22(4):413–434.

Jackman, Simon. 2001. "Multidimensional analysis of roll call data via Bayesian simulation: identification, estimation, inference, and model checking." *Political Analysis* 9:227–241.

Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. Wiley.

Junker, Brian, Lynne Steuerle Schofield and Lowell J. Taylor. 2012. "The use of cognitive ability measures as explanatory variables in regression analysis." *IZA Journal of Labor Economics* 1(1996):1–19.

Lau, Richard R. and David P. Redlawsk. 1997. "Voting Correctly." *American Political Science Review* 91(3):585–598.

Lau, Richard R., Parina Patel, Dalia F. Fahmy and Robert R. Kaufman. 2014. "Correct Voting Across Thirty-Three Democracies: A Preliminary Analysis." *British Journal of Political Science* 44(2):239–259.

Levendusky, Matthew S. and Simon Jackman. 2003. "Reconsidering the Measurement of Political Knowledge." Working Paper.

Poole, Keith T. and Howard. Rosenthal. 2000. *Congress : A Political-Economic History of Roll Call Voting*. Oxford University Press.

Rivers, Douglas. 2003. "Identification of Multidimensional Spatial Voting Models." Typescript, Stanford University.