Supplementary Material

Jørgen Bølstad

This material belongs to the article "Capturing Rationalization Bias and Differential Item Functioning: A Unified Bayesian Scaling Approach" in *Political Analysis*.

Contents

1 Question Wording

The survey data are from the European Election Study 2009 (EES 2011), where the question regarding self-placement was phrased this way:

> In political matters people talk of ‘the left’ and ‘the right’. What is your position? Please indicate your views using any number on a scale from 0 to 10, where 0 means ‘left’ and 10 means ‘right’. Which number best describes your position?

Similar questions were asked regarding party positions, rotating the order of the parties across respondents. Respondents were also asked to report their party preferences, expressed as the probability of ever voting for a party:

> We have a number of parties in [this country] each of which would like to get your vote. How probable is it that you will ever vote for the following parties? Please specify your views on a scale where 0 means ‘not at all probable’ and 10 means ‘very probable’.

2 Additional Simulation Results

This section reports results for additional test criteria and scenarios. One of the new criteria tests whether the models reveal the true extent of proximity voting: I first calculate the absolute distances between stimuli and voters in both the true and the rescaled data. Then, for each individual, I correlate these sets of distances with their preferences to obtain measures of true and estimated proximity voting. Based on these measures, I report the average bias and root mean squared error (RMSE) across all individuals for each dataset. For the ISR model, I also check how well it estimates the γ-parameters by reporting the average bias and RMSE across all individuals for each dataset.[1]

Figure S1 reports on the same analyses as in the article, but includes the additional test criteria. The two upper panels are also shown in the article. The two lower panels in the left column of the figure show whether the models succeed in estimating the extent of genuine proximity voting. As we would expect, in the scenario with no rationalization, all models yield unbiased estimates. However, as the degree of rationalization increases, analyses based on AM-type models increasingly overestimate the extent of proximity voting. This also results in higher RMSEs for these estimates, while the ISR model's estimates have more stable RMSEs. On these two test criteria, the differences between the BAM2 model and the other AM-type models are modest. Finally, the two lower panels in the right column of Figure S1 show the extent to which the ISR model succeeds at estimating the degree of rationalization. The RMSEs are relatively low considering the scale, and the estimates are essentially unbiased in all scenarios.

Figure S2 reports the same test criteria for data with larger errors (average standard deviations around 1.5). The results generally show the same patterns as before, although the BAM2 model now clearly outperforms the older AM-type models even in the absence of rationalization: Its more informative priors make it considerably more robust to noise. In terms of uncovering voter positions, the BAM2 model's performance is closer to the ISR model than the other AM-type models in these scenarios. However, the ISR model still performs somewhat better (and shows more stable performance) also in this regard.

[1]Given the bounded nature of the γ-parameters, I use the posterior median to obtain point estimates, except in the scenario with zero rationalization: As the true values here lie on boundary of the parameter space, I estimate posterior modes, after reflecting the distributions at the bounds.

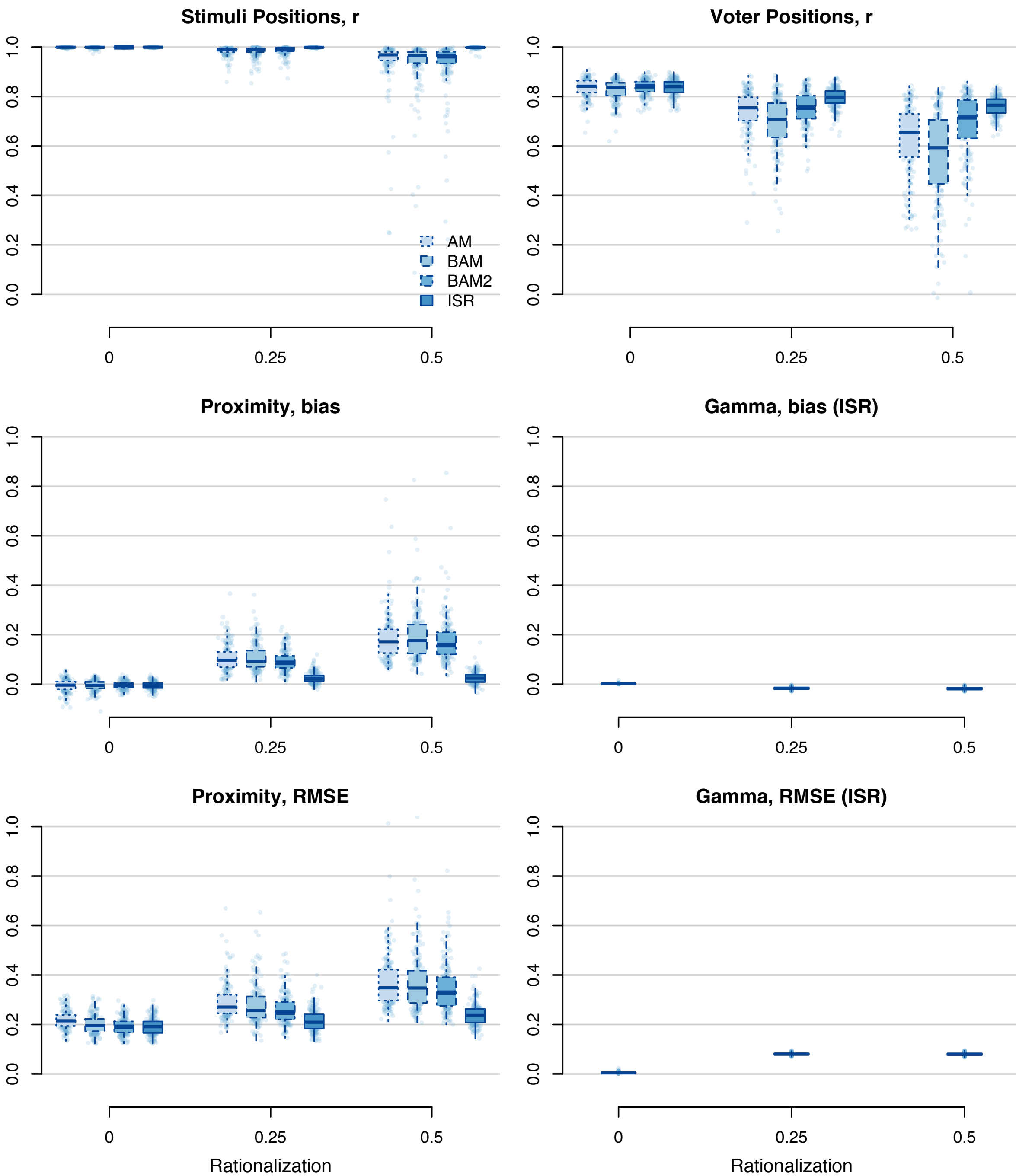

Figure S1: Model performance in simulations with small errors.

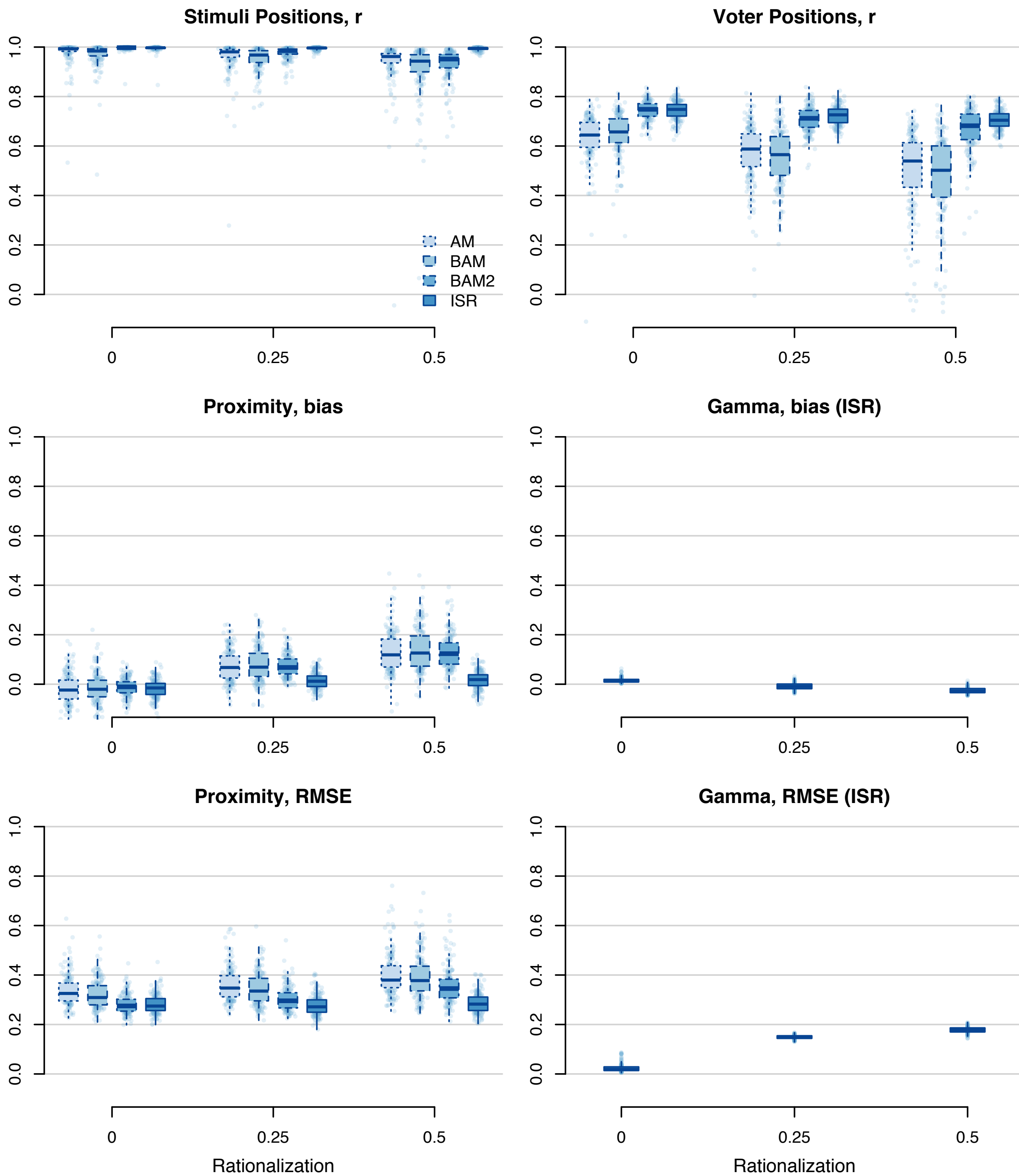

Figure S2: Model performance in simulations with large errors.

3 Plots Including BAM2

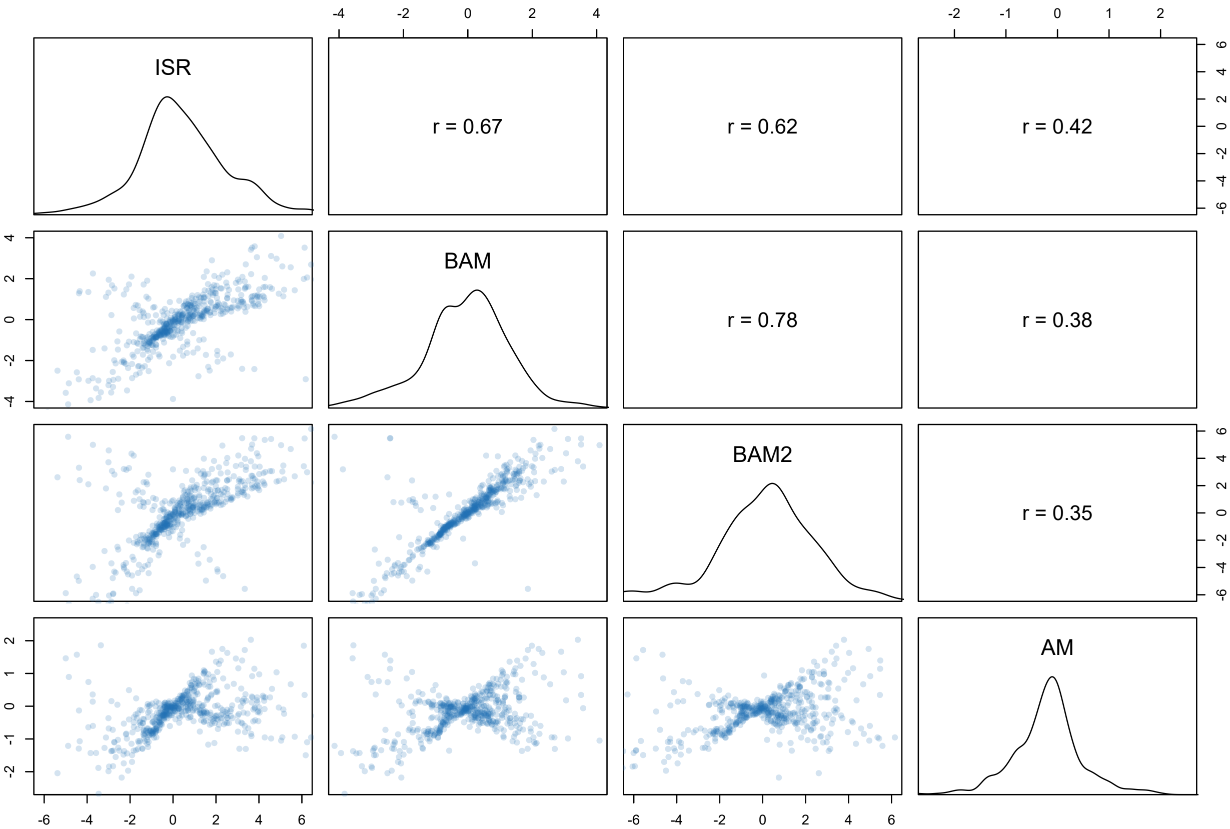

Figure S3: Correlations estimated voter positions in the UK including BAM2.

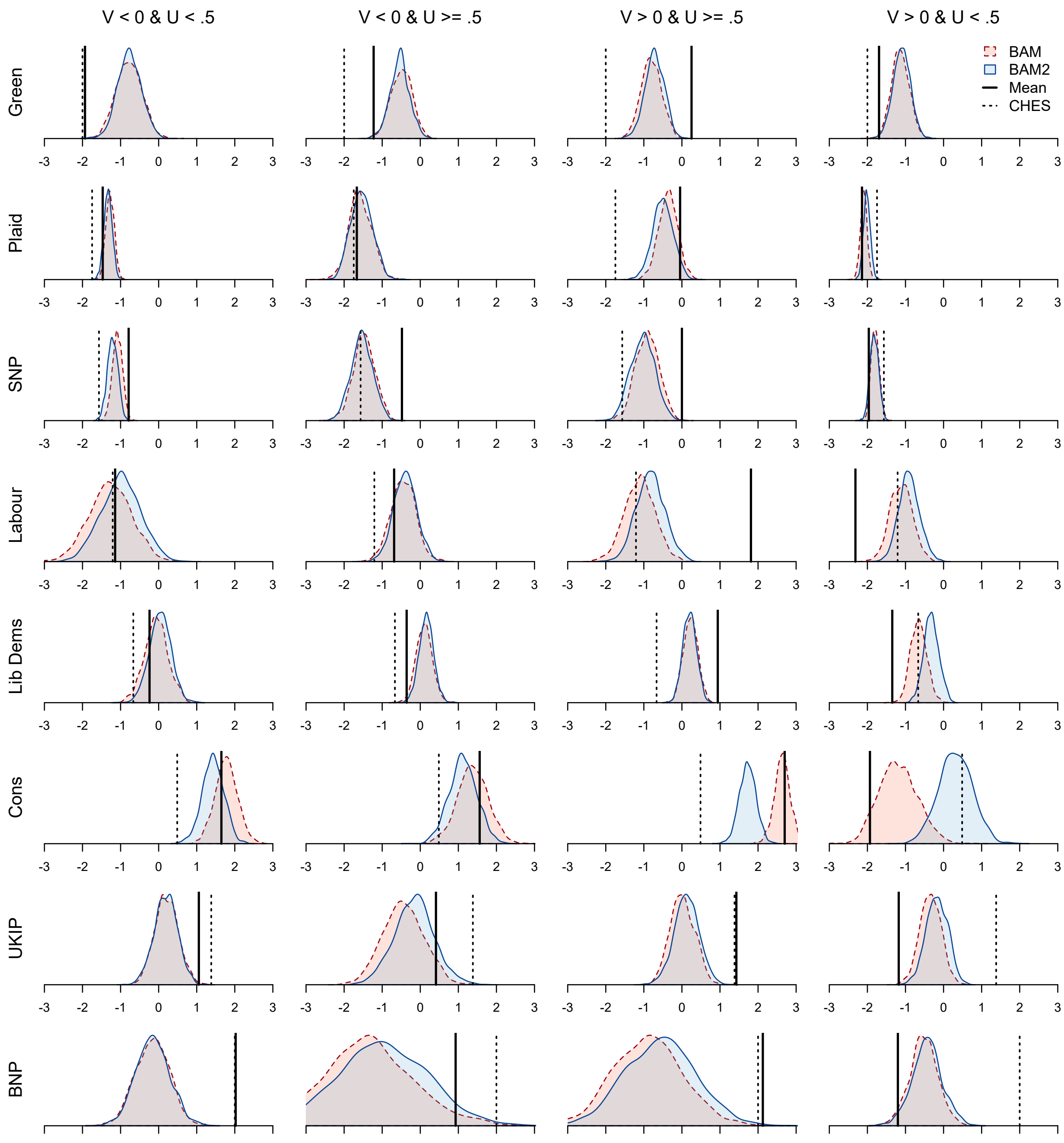

Figure S4: Posterior predictive densities for the BAM and BAM2 models in the UK.

4 Traceplots

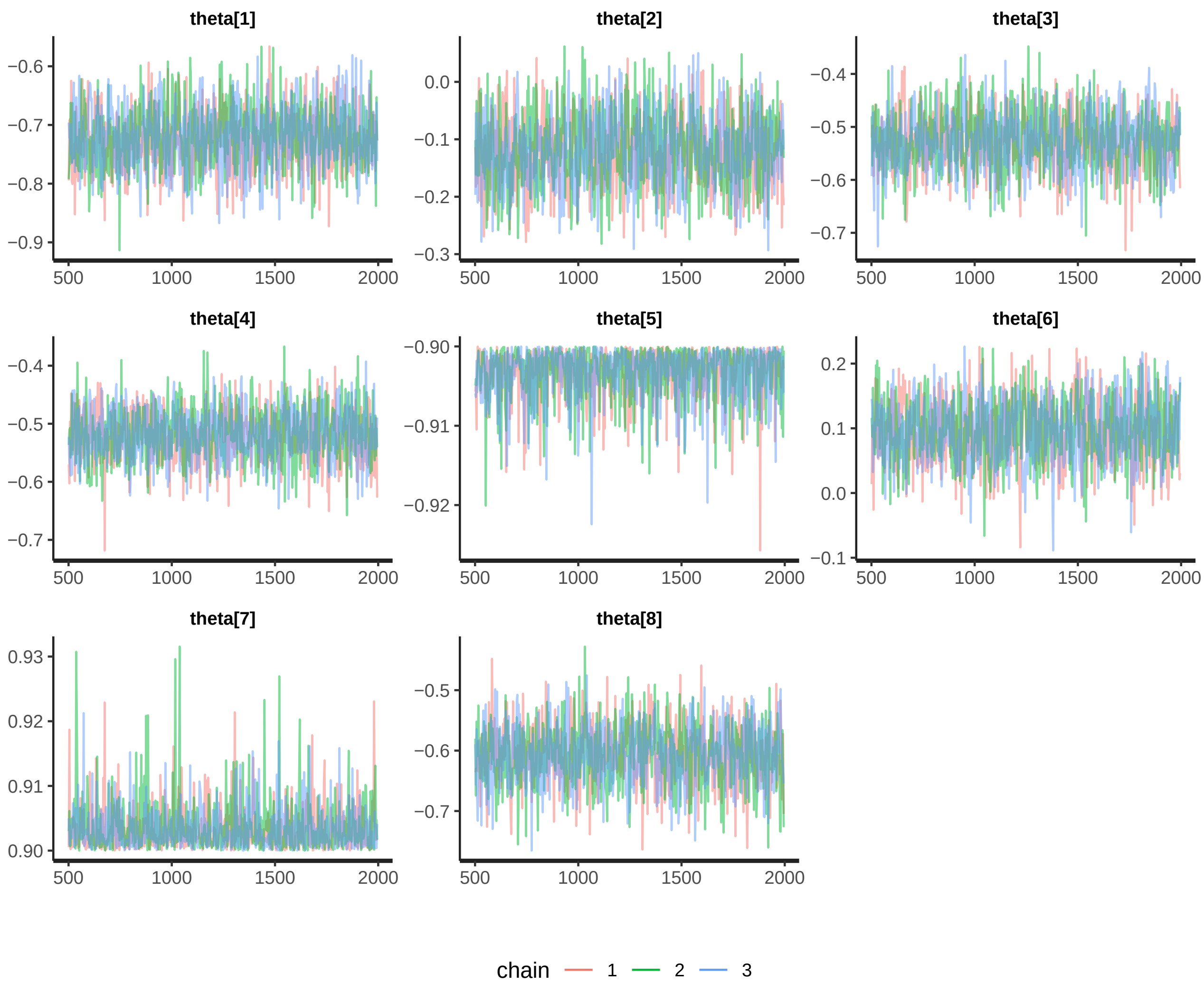

Figure S5: Traceplots of θ for the AM* model in the UK.

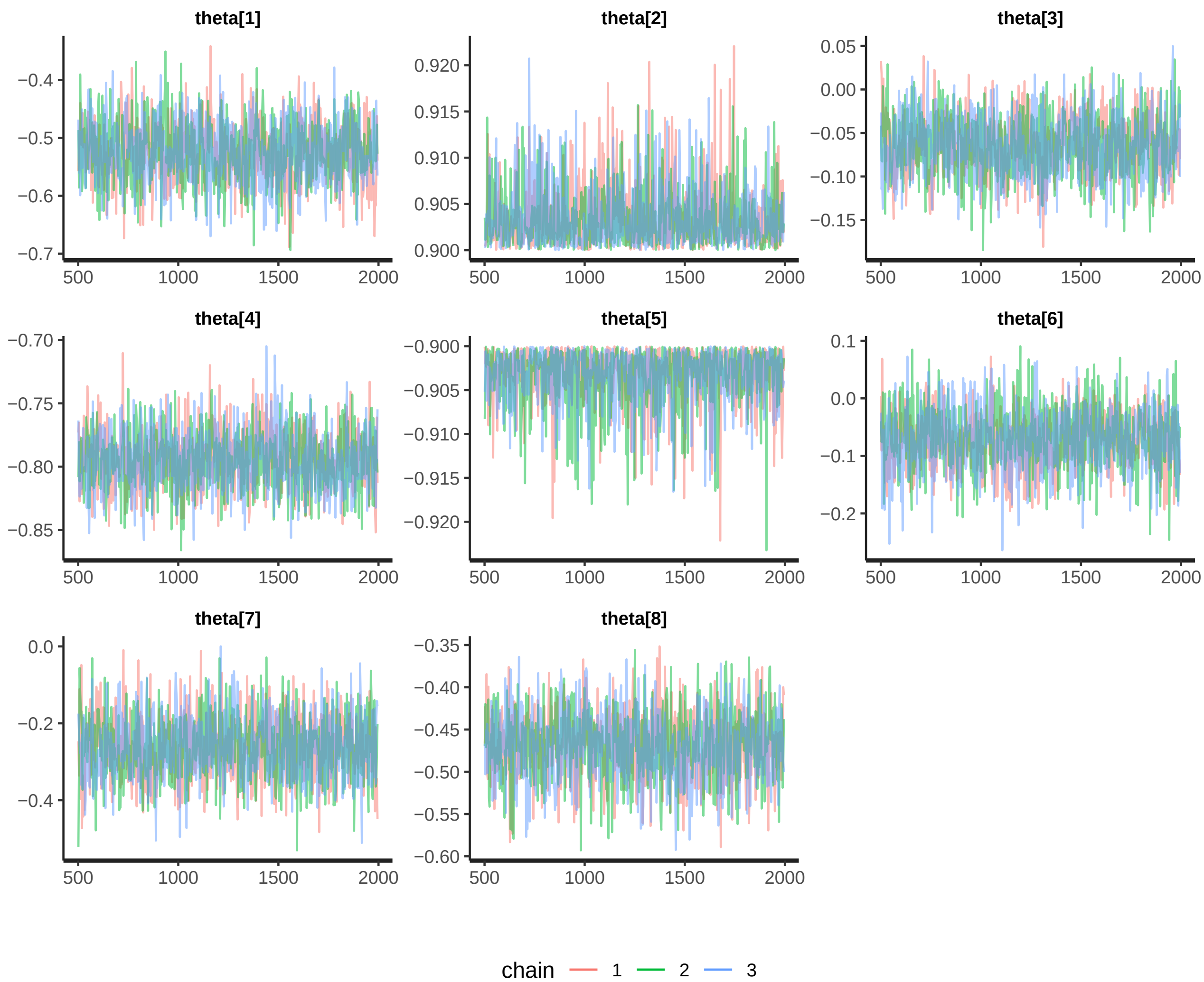

Figure S6: Traceplots of θ for the BAM model in the UK.

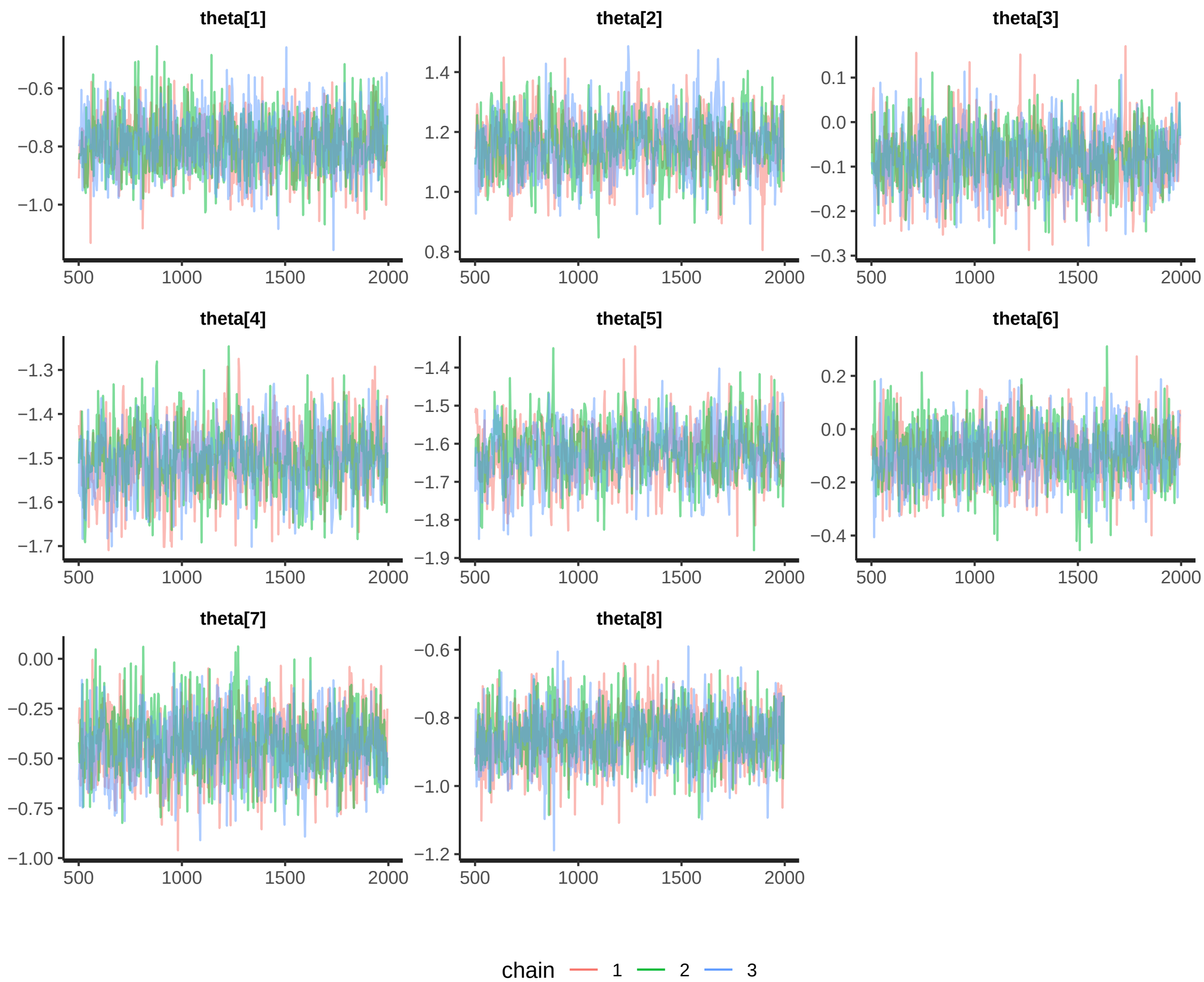

Figure S7: Traceplots of θ for the BAM2 model in the UK.

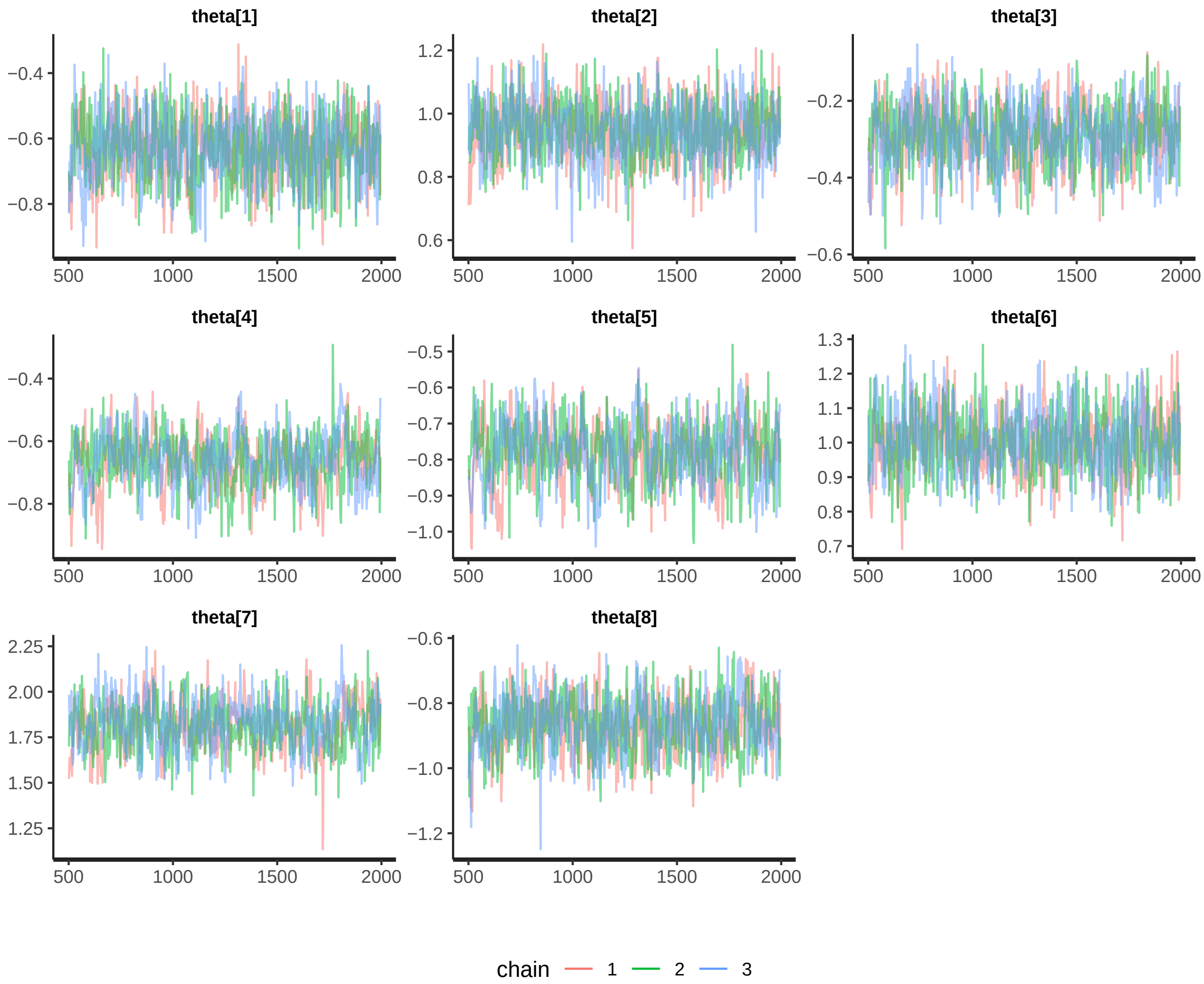

Figure S8: Traceplots of θ for the ISR model in the UK.

5 Results for 14 Additional Countries

To show a more general picture of how the models perform in a variety of settings, this section reports results for the AM*, BAM, BAM2, and ISR model in 14 additional countries.[2] For Bulgaria, a homoskedastic version of the ISR model is used to ensure efficient sampling and convergence to the target distribution.[3] As shown in column 5 of Table S1, the observed sample mean placements are strongly correlated with the CHES expert placements in most of these cases, which suggests that the models have a relatively easy task in recovering the party positions.[4] Indeed, all models tend to perform very well in this regard: The median correlation is around .94 and .95 for each of them. These results are consistent with the Monte Carlo simulations reported in this study, showing that the AM-type models generally perform well in terms of recovering stimuli positions, even in the face of rationalization.

However, Table S2 further reports the Watanabe-Akaike information criterion (WAIC), which suggests that the ISR model fits the data significantly better in nearly every single case (Watanabe 2010).[5] This suggests that a notable degree of rationalization is present in all of these countries. To assess the extent of rationalization estimated by the ISR model, column 4 of Table S1 reports the average of the γ-parameters for each country. This average typically lies between .3 and .4, pointing to a significant degree of rationalization in all of these countries. This is important, because the Monte Carlo simulations show that AM-type models tend to give biased voter positions estimates under such circumstances.

To assess whether the models do indeed produce different voter position estimates, the last three columns of Table S3 shows pairwise correlations between estimates from each of the models. Interestingly, the AM* and BAM models tend to give somewhat different estimates, having a median correlation of .87 between their estimates. This is worth noting, as the choice between the two error specifications may be more consequential than users of these models appreciate. As we would expect, the AM-type models also produce significantly

[2]Four countries were excluded by Lo, Proksch and Gschwend (2014) due to coding issues in the EES data, and they are also excluded here: Belgium, Denmark, Sweden, and Spain. The following countries are also left aside, because they have five or fewer party observations available to validate the party position estimates: Czech Republic, Cyprus, Germany, Latvia, Luxembourg, Malta, Poland, and Portugal. At the individual level, I require that respondents report a self-placement and use at least three unique party placements, while I permit missing preferences and placements for up to one party.

[3]This case could possibly also be handled by running the chains longer, but I rather use this as an opportunity to introduce a simplified version of the model, which may prove useful also in other applications.

[4]In their analysis of American data, Hare et al. (2015) report a correlation of 1.00 between BAM estimates and mean placements.

[5]The exception is Bulgaria, where the BAM2 model has a slighly lower WAIC.

Table S1: Summary of results for 14 countries.

			Mean	r, CHES				
	N	J	Rat.	Means	AM*	BAM	BAM2	ISR
Austria	694	8	0.38	0.99	0.98	0.98	0.99	0.99
Bulgaria	380	8	0.32	0.95	0.96	0.96	0.95	0.96
Estonia	502	6	0.33	0.93	0.85	0.86	0.92	0.94
Greece	824	6	0.35	0.98	0.99	0.99	0.99	0.99
Finland	865	8	0.27	0.97	0.97	0.98	0.98	0.98
France	767	8	0.42	0.92	0.94	0.95	0.94	0.94
Hungary	639	7	0.39	0.92	0.93	0.94	0.93	0.94
Ireland	779	6	0.46	0.97	0.91	0.94	0.97	0.97
Italy	586	8	0.33	0.99	0.99	0.99	0.99	0.99
Lithuania	358	10	0.29	0.63	0.56	0.58	0.61	0.58
Netherlands	813	11	0.34	0.98	0.97	0.97	0.98	0.98
Romania	355	7	0.48	0.47	0.58	0.67	0.55	0.69
Slovenia	718	9	0.29	0.88	0.87	0.87	0.88	0.88
Slovakia	618	8	0.29	0.83	0.79	0.81	0.78	0.81

Note: The columns under "r, CHES" report pairwise correlations between CHES expert placements and the specified type of party position estimates. "Means" refers to the sample mean party placements.

Table S2: Watanabe-Akaike information criterion (WAIC).

	AM*	BAM	BAM2	ISR
Austria	24502.41	21792.20	21951.17	20799.05
Bulgaria	13659.74	13176.42	12404.62	12664.18
Estonia	13248.08	12564.42	12186.70	11581.17
Greece	28626.88	26545.65	25805.43	24423.98
Finland	25135.52	22762.58	22265.06	19890.13
France	22052.53	20431.67	19875.37	18529.24
Hungary	18929.11	18180.73	17761.42	16135.17
Ireland	23095.29	20946.38	21239.74	19403.62
Italy	19193.63	17291.58	17027.12	15606.85
Lithuania	15410.87	14958.38	14606.05	14022.69
Netherlands	37180.20	34313.68	34019.66	31939.56
Romania	13007.49	11726.98	12343.32	11196.51
Slovenia	22602.03	21253.85	20768.83	20116.30
Slovakia	29078.17	27403.38	26928.23	26037.77

different estimates from the ISR model. The median correlation with the ISR estimates are .76 for the AM* model, .73 for the BAM model, and .82 for the BAM2 model. This illustrates once again that the models can give notably different results, and that the choice of model can be important even when the stimuli position estimates are very similar and appear highly valid (such as in Austria).

Table S3: Pairwise correlations of voter position estimates.

Model 1 Model 2	AM* BAM	AM* BAM2	BAM BAM2	AM* ISR	BAM ISR	BAM2 ISR
Austria	0.65	0.79	0.78	0.73	0.53	0.75
Bulgaria	0.81	0.91	0.73	0.90	0.88	0.84
Estonia	0.73	0.72	0.76	0.66	0.66	0.89
Greece	0.96	0.88	0.90	0.87	0.89	0.92
Finland	0.82	0.74	0.90	0.77	0.74	0.74
France	0.96	0.93	0.91	0.83	0.82	0.83
Hungary	0.96	0.85	0.87	0.80	0.77	0.84
Ireland	0.84	0.76	0.66	0.67	0.58	0.81
Italy	0.95	0.78	0.84	0.73	0.71	0.78
Lithuania	0.93	0.79	0.77	0.61	0.63	0.78
Netherlands	0.90	0.88	0.90	0.78	0.83	0.85
Romania	0.59	0.57	0.54	0.56	0.49	0.74
Slovenia	0.77	0.87	0.86	0.74	0.69	0.80
Slovakia	0.92	0.86	0.83	0.84	0.83	0.94

6 Model Codes

Stan Code for the AM* Model

```
data {
  int<lower = 1> N;                        // n of individuals
  int<lower = 1> J;                        // n of items
  int<lower = 1, upper = J> L;             // left pole
  int<lower = 1, upper = J> R;             // right pole
  real Y[N, J];                            // reported stimuli position
  int<lower = 0, upper = 1> M[N, J];       // indicator of missing values
}

parameters {
  real alpha[N];                           // shift parameter
  real beta[N];                            // stretch parameter
  real<lower = -1.1, upper = -.9> thetal; // left pole stimuli
  real<lower = .9, upper = 1.1> thetar;   // right pole stimuli
  real thetam[J];                          // remaining stimuli
  real<lower = 0> tau;                     // homoskedastic precision
}

transformed parameters {
  real theta[J];                           // latent stimuli position
  matrix[N, J] log_lik;                    // pointwise log-likelihood
  real sigma = sqrt(1 / tau);
  theta = thetam;
  theta[L] = thetal;
  theta[R] = thetar;
  for (i in 1:N) {
    for (j in 1:J) {
      if (M[i, j] == 0) {
        log_lik[i, j] = normal_lpdf(Y[i, j] | alpha[i] + beta[i] * theta[j], sigma);
      }
      else {log_lik[i, j] = 0;}
    }
  }
}

model {
  alpha ~ uniform(-100, 100);
  beta ~ uniform(-100, 100);
  thetam ~ normal(0, 1);
  thetal ~ normal(0, 1);
  thetar ~ normal(0, 1);
  tau ~ gamma(.1, .1);

  target += sum(log_lik);
}
```

Stan Code for the BAM Model

```
// Adapted from the JAGS code by Hare et al. (AJPS 2015).

data {
  int<lower = 1> N;                        // n of individuals
  int<lower = 1> J;                        // n of items
  int<lower = 1, upper = J> L;             // left pole
  int<lower = 1, upper = J> R;             // right pole
  real Y[N, J];                            // reported stimuli position
  int<lower = 0, upper = 1> M[N, J];       // indicator of missing values
}

parameters {
  real alpha[N];                           // shift parameter
  real beta[N];                            // stretch parameter
  real<lower = -1.1, upper = -.9> thetal; // left pole stimuli
  real<lower = .9, upper = 1.1> thetar;   // right pole stimuli
  real thetam[J];                          // remaining stimuli
  real<lower = 0> nu;                      // hyperparameter
  real<lower = 0> omega;                   // hyperparameter
  vector<lower = 0>[N] tau;                // individual precision
  row_vector<lower = 0>[J] eta;            // stimuli precision
}

transformed parameters {
  real theta[J];                           // latent stimuli position
  matrix[N, J] log_lik;                    // pointwise log-likelihood
  matrix[N, J] sigma;
  theta = thetam;
  theta[L] = thetal;
  theta[R] = thetar;
  sigma = sqrt(1 ./ (tau * eta) );
  for (i in 1:N) {
    for (j in 1:J) {
      if (M[i, j] == 0) {
        log_lik[i, j] = normal_lpdf(Y[i, j] | alpha[i] + beta[i] * theta[j], sigma[i, j]);}
      else {log_lik[i, j] = 0;}
    }
  }
}

model {
  alpha ~ uniform(-100, 100);
  beta ~ uniform(-100, 100);
  thetam ~ normal(0, 1);
  thetal ~ normal(0, 1);
  thetar ~ normal(0, 1);
  nu ~ gamma(0.1, 0.1);
  omega ~ gamma(0.1, 0.1);
  tau ~ gamma(nu, omega);
  eta ~ gamma(0.1, 0.1);
  target += sum(log_lik);
}
```

Stan Code for the BAM2 Model

```
data {
  int<lower = 1> N;                        // n of individuals
  int<lower = 1> J;                        // n of items
  int<lower = 1> B;                        // scale bound
  int<lower = -B, upper = B> Y[N, J];      // reported stimuli position
  int<lower = 0, upper = 1> M[N, J];       // indicator of missing values
}

parameters {
  vector[N] alpha_raw;                     // shift parameter, raw
  real beta[N];                            // stretch parameter
  real theta[J];                           // latent stimuli position
  real<lower = 0> lambda;                  // sd of alpha
  real<lower = 1> nu;                      // hyperparameter
  real<lower = 0> tau;                     // hyperparameter
  vector<lower = 0>[N] eta;                // mean ind. variance x J^2
  simplex[J] rho;                          // stimuli-shares of variance
}

transformed parameters {
  vector[N] alpha;                         // shift parameter
  matrix<lower = 0>[N, J] sigma;           // sd's of errors
  matrix[N, J] log_lik;                    // pointwise log-likelihood
  alpha = alpha_raw * lambda;
  sigma = sqrt(eta) * to_row_vector(rho);
  for (i in 1:N) {
    for (j in 1:J) {
      if (M[i, j] == 0) {
        log_lik[i, j] = normal_lpdf(Y[i, j] | alpha[i] + beta[i] * theta[j], sigma[i, j]);
      }
      else {log_lik[i, j] = 0;}
    }
  }
}

model {
  alpha_raw ~ normal(0, 1);
  lambda ~ cauchy(0, B);
  beta ~ normal(1, 1);
  theta ~ normal(0, 10);
  nu ~ cauchy(0, 50);
  tau ~ cauchy(0, J * B);
  eta ~ scaled_inv_chi_square(nu, tau);
  rho ~ dirichlet(rep_vector(5, J));

  target += sum(log_lik);
}
```

Stan Code for the ISR Model

```
data {
  int<lower = 1> N;                         // n of individuals
  int<lower = 1> J;                         // n of items
  int<lower = 1> B;                         // scale bound
  int<lower = -B, upper = B> Y[N, J];       // reported stimuli position
  int<lower = -B, upper = B> V[N];          // reported voter position
  real<lower = 0, upper = 1> U[N, J];       // reported voter preference
  int<lower = 0, upper = 1> M[N, J];        // indicator of missing values
}

transformed data {
  real<lower = -B, upper = B> p1[N, J];
  real<lower = -B, upper = B> p2[N, J];
  real z;
  z = (5.0 / (B * 1.0));
  for (i in 1:N) {
    for (j in 1:J) {
      p1[i, j] = U[i, j] * V[i] + B * U[i, j] - B;
      p2[i, j] = U[i, j] * V[i] - B * U[i, j] + B;
    }
  }
}

parameters {
  vector[N] alpha_raw;                      // shift parameter, raw
  real beta[N];                             // stretch parameter
  real theta[J];                            // latent stimuli position
  real<lower = 0> lambda;                   // sd of alpha
  real<lower = 1> nu;                       // hyperparameter
  real<lower = 0> tau;                      // hyperparameter
  vector<lower = 0>[N] eta;                 // mean ind. variance x J^2
  simplex[J] rho;                           // stimuli-shares of variance
  real<lower = 0, upper = 1> gamma[N];      // rationalization
  real<lower = 1> gam_a;                    // hyperparameter
  real<lower = 1> gam_b;                    // hyperparameter
  real<lower = 0, upper = 1> delta;         // weight in mixing prop.
}

transformed parameters {
  vector[N] alpha;                          // shift parameter
  real mu0;                                 // dif-adjusted mean
  matrix<lower = 0>[N, J] sigma;            // sd's of errors
  matrix[N, J] log_lik;                     // pointwise log-likelihood
  alpha = alpha_raw * lambda;
  sigma = sqrt(eta) * to_row_vector(rho);
  for (i in 1:N) {
    for (j in 1:J) {
      if (M[i, j] == 0) {
        mu0 = alpha[i] + beta[i] * theta[j];
        log_lik[i, j] = log_mix( (.5 * (1 - delta)) + (delta * inv_logit(z * (V[i] - mu0))),
          normal_lpdf(Y[i, j] | (1 - gamma[i]) * mu0 + gamma[i] * p1[i, j], sigma[i, j]),
          normal_lpdf(Y[i, j] | (1 - gamma[i]) * mu0 + gamma[i] * p2[i, j], sigma[i, j]) );
```

```
      }
      else {log_lik[i, j] = 0;}
    }
  }
}

model {
  alpha_raw ~ normal(0, 1);
  lambda ~ cauchy(0, B);
  beta ~ normal(1, 1);
  gamma ~ beta(gam_a, gam_b);
  gam_a ~ gamma(1.5, .5);
  gam_b ~ gamma(1.5, .5);
  delta ~ beta(3, 1.1);
  theta ~ normal(0, 10);
  nu ~ cauchy(0, 50);
  tau ~ cauchy(0, J * B);
  eta ~ scaled_inv_chi_square(nu, tau);
  rho ~ dirichlet(rep_vector(5, J));

  target += sum(log_lik);
}
```

References

EES. 2011. "European Parliament Election Study 2009, Voter Study, Advance Release, 23/06/2011." http://www.piredeu.eu.

Hare, Christopher, David A. Armstrong, Ryan Bakker, Royce Carroll and Keith T. Poole. 2015. "Using Bayesian Aldrich-McKelvey Scaling to Study Citizens' Ideological Preferences and Perceptions." *American Journal of Political Science* 59(3):759–774.

Lo, James, Sven-Oliver Proksch and Thomas Gschwend. 2014. "A Common Left-Right Scale for Voters and Parties in Europe." *Political Analysis* 22(2):205–223.

Watanabe, Sumio. 2010. "Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory." *Journal of Machine Learning Research* 11:3571–3594.