# Supplemental Files for "Measuring Foreign Policy Positions of Members of the US Congress"

**Table of Content**

**R code for Dynamic IRT estimation for the Senate Data**

```
library(MCMCpack)
out <- MCMCdynamicIRT1d(Data, # "Data" = Roll Call Data
           item.time.map=time, # time = 1, 2, ..., T
           mcmc=50000, burnin=5000, thin=100,
           theta.constraints=list(A9369A="+",      # Strom Thurmond
                                  A14105A="+",     # Jesse Helms
                                  A14920A="-",     # John Kerry
                                  A10808A="-"))    # Edward Kennedy
           ### Note "A9369A" is the name of rows in "Data"
           #   that correspond to Strom Thurmond
           ### "Data" should include rowname, time, and voting records.
```

## Monte Carlo Analysis - Interest Group Ratings and Extremism

The purpose of this Monte Carlo analysis is to provide evidence that interest group ratings overemphasize extremism and that IRT models do not underemphasize polarization. In the experiment, I generated 434 ideal points (258 Democrats and 176 Republicans, as in the 103rd House) and 100 cutpoints (there were 108 non-unanimous foreign policy votes in the 103rd House). I simulated the varying degrees of ideological polarization by generating ideal points from three different scenarios. First, I generated all the ideal points from a standard normal distribution (a situation with no polarization). Second, I generated Democrats from $\mathcal{N}(-1, 1)$ and Republicans from $\mathcal{N}(1, 1)$. Finally, to simulate a polarized Congress, I generated Democrats from $\mathcal{N}(-2, 1)$ and Republicans from $\mathcal{N}(2, 1)$. In all three cases, the cutpoints are generated from a standard normal distribution.

Figure 1 below shows the performance of the IRT model in all three situations. In each panel, the horizontal dimension represents true ideal points, while the vertical dimension represents the IRT estimates of ideal points. In general, the IRT model performs well, indicated by the dots along the 45 degree lines. However, in the polarized situation (the right panel), the IRT model was less successful in capturing the true ideal points of extreme members. This can be explained by the fact that the cutpoints were generated from a standard normal distribution, while ideal points were generated from $\mathcal{N}(-2, 1)$ and $\mathcal{N}(2, 1)$. That is, while almost all the cutpoints ranged between -3 and 3, ideal points ranged between -5 and 5. Simply put, there were not enough cutpoints that divided extreme members on both sides, causing many extreme members to vote identically. This is an issue with the data rather than the model. Nevertheless, this data problem should be rare in reality. As the top panel of Figure 3 shows, there were still a dozen votes that divided extreme members (with a majority size greater than 80%) even in a polarized legislature, like the 103rd House. In fact, when I generated the cutpoints from $\mathcal{N}(0, 2)$, extreme members' ideal points were captured a lot better. All in all, the main point of this experiment is clear. That is, there is no evidence that the IRT model systematically underemphasizes polarization.

Next, to simulate interest group ratings, I follow the procedure of Americans for Democratic Actions (ADA) rankings. Every year, ADA selects 20 important votes and computes the percentage that a member voted with ADA. Thus, a member who voted with ADA in all 20 votes gets a score of 100%, while a member who did not vote with ADA on any of the 20 votes gets 0%. Here, I operationalize the selection of "important" votes by looking at the size of the majority (%). This is because interest groups ratings tend to pick controversial votes that divide political parties and exclude less controversial votes that have bipartisan support. For instance, the 40 House votes selected by ADA in 1993 and 1994 (103rd House) had an average vote margin of 9%. That is, the average size of the majority was 59%. Figure 2 illustrates ADA's tendency to focus on close votes in comparison to the majority size of all foreign policy votes in the 103rd House. It shows that ADA selected only one vote that had a majority size greater than 80%, while there were more than a dozen foreign policy votes that had a majority size greater than 80%. Thus, to examine how this tendency to focus on controversial votes affects interest group ratings, I varied the size of majority on the votes selected by ADA (60% or less, 70% or less, and 80% or less). If the tendency to focus on controversial votes causes overemphasis of extremism in interest group ratings, we should observe that this problem gets worse as the maximum majority size gets smaller (i.e. when it is 60% rather than 80%).

Figure 3 shows the results. The panels in the first row show the results for the simulated ADA scores that used votes with a majority size of 60% or less. In the second row, ADA scores are simulated with votes that have a majority size of 70% or less. In the third row, ADA scores are simulated with votes that have a majority size of 80% or less. The degree of partisan polarization varies across columns. The first column simulates no polarization situation, while the third column simulates a stronger degree of polarization. The results show that the simulated ADA score performs better when the votes are not confined to a small number of controversial votes, as shown in the bottom panels. When ADA selects only close votes, as shown in the top panels, the ADA scores' tendency to overemphasize polarization becomes severe. This confirms that the practice of using a small number of controversial votes to compute interest group ratings is problematic as it overemphasizes the degree of polarization. This problem is mitigated when the actual degree of party polarization is high as shown in the panels in the third column. Even under these conditions, the simulated ADA scores still tend to overemphasize the degree of polarization.

To check if the above results are due to simulation errors, I repeated the simulations 50 times. Figure 4 reports the results of two polarization scenarios (DEM $\sim \mathcal{N}(-1,1)$ and GOP $\sim \mathcal{N}(1,1)$ vs. DEM $\sim \mathcal{N}(-2,1)$ and GOP $\sim \mathcal{N}(2,1)$) and two different majority sizes used by ADA (60 % or less vs. 80% or less). The results confirm that the above findings are not due to simulation errors.

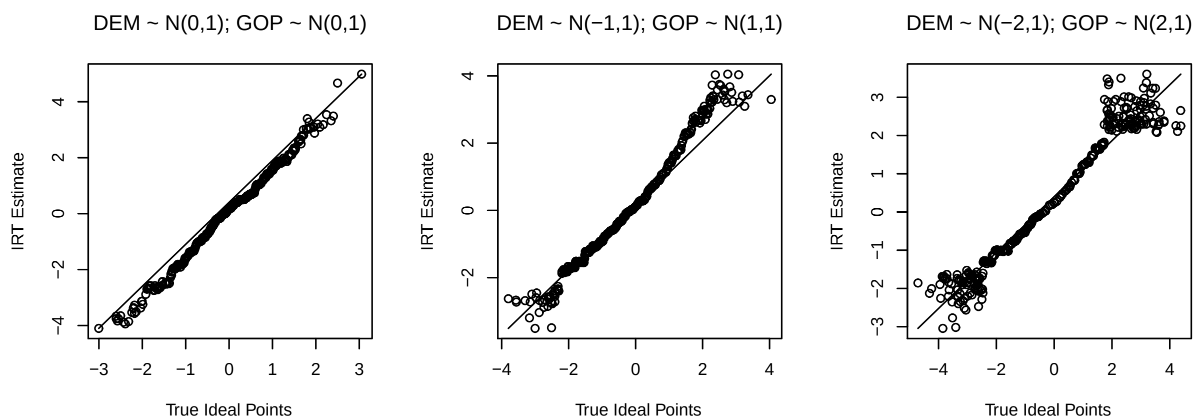



Figure 1: **Monte Carlo Analysis of the Performance of IRT Models in Different Party Polarization Situations.** In the left panel, all ideal points are generated from a standard normal distribution. In the center panel, Democrats (258) are generated from $\mathcal{N}(-1, 1)$ and Republicans (176) from $\mathcal{N}(1, 1)$. In the right panel, Democrats are generated from $\mathcal{N}(-2, 1)$ and Republicans from $\mathcal{N}(2, 1)$. All cutpoints are generated from a standard normal distribution in all three cases.

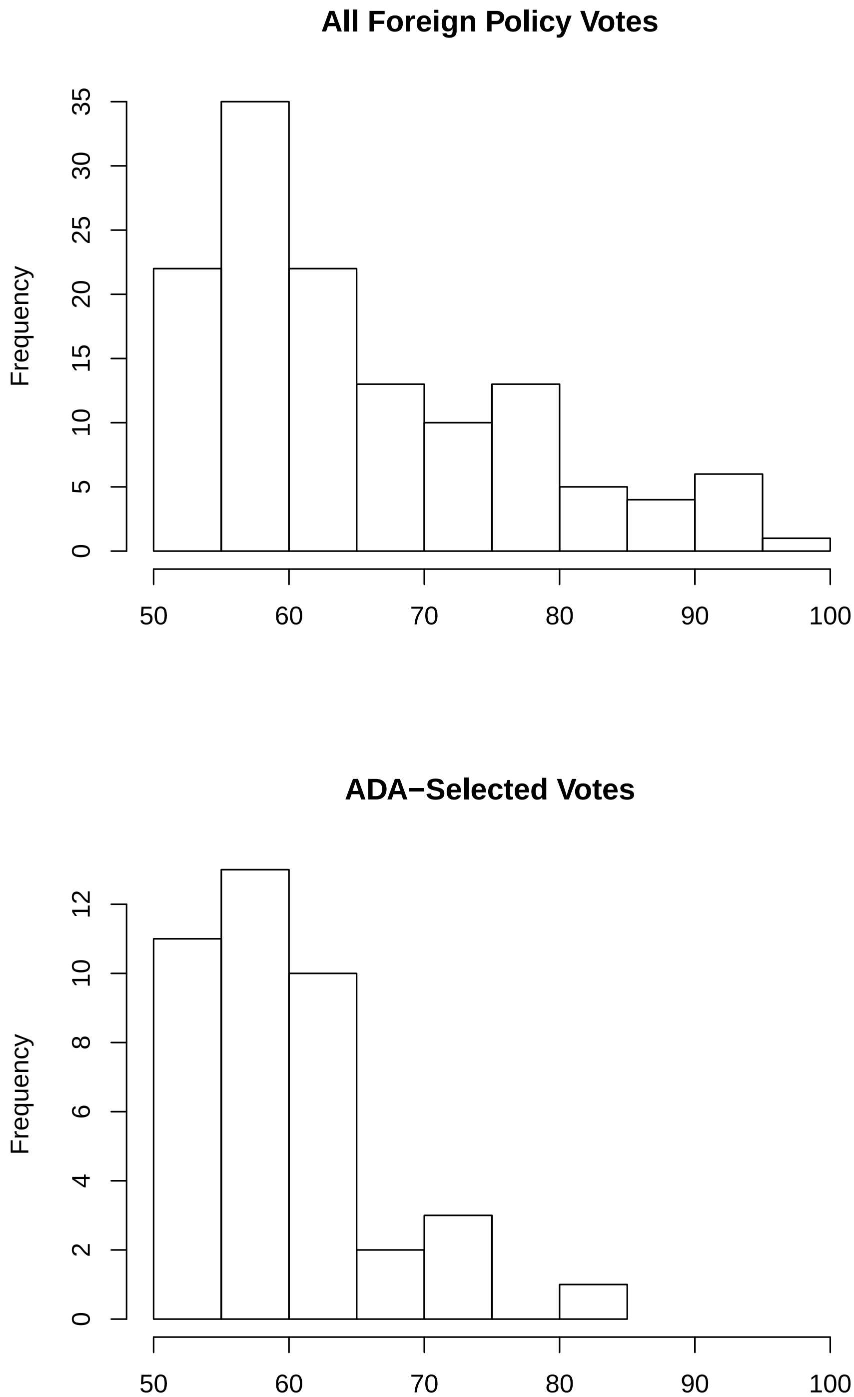



Figure 2: **The Majority Size (%) of Foreign Policy Votes and ADA-Selected Votes (103rd House).**

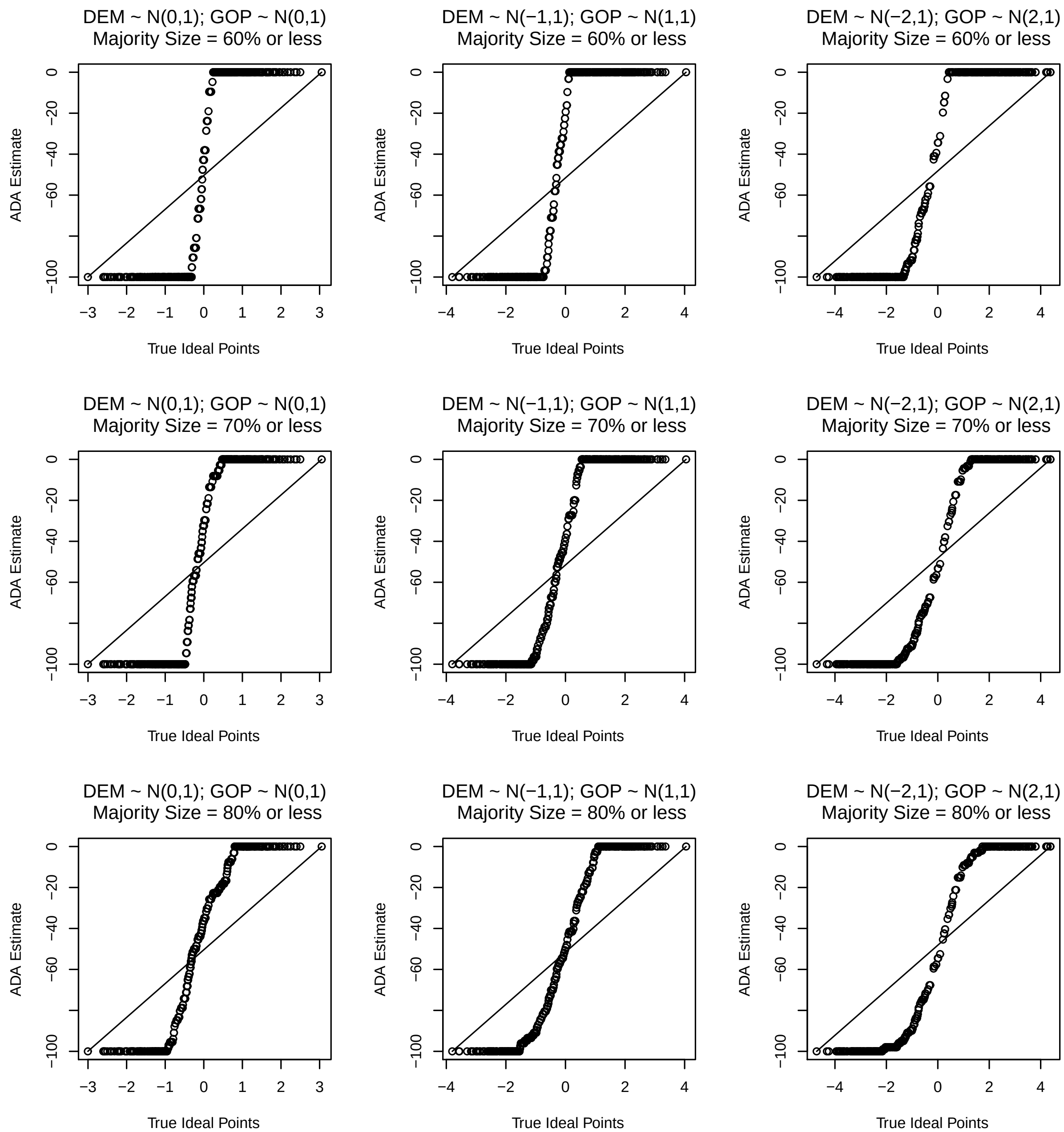



Figure 3: **Monte Carlo Analysis of Interest Group Ratings' Extremism.** The degree of polarization varies as in Figure 1. The majority size of votes selected by ADA varies across rows. In the first row, ADA ratings are generated using simulated votes that have a majority size of 60% or less; the second row, 70% or less; and the third rows, 80% or less.

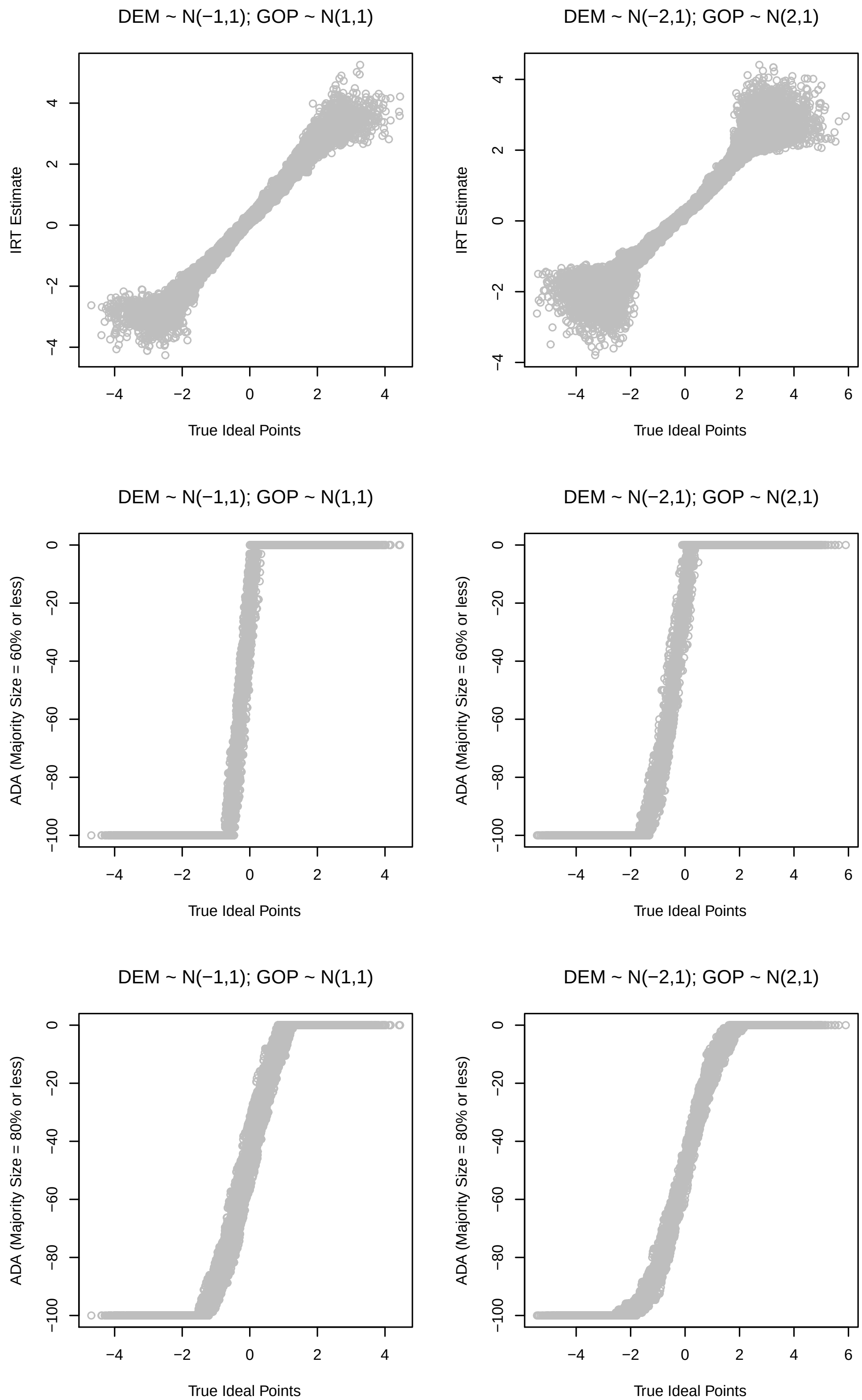
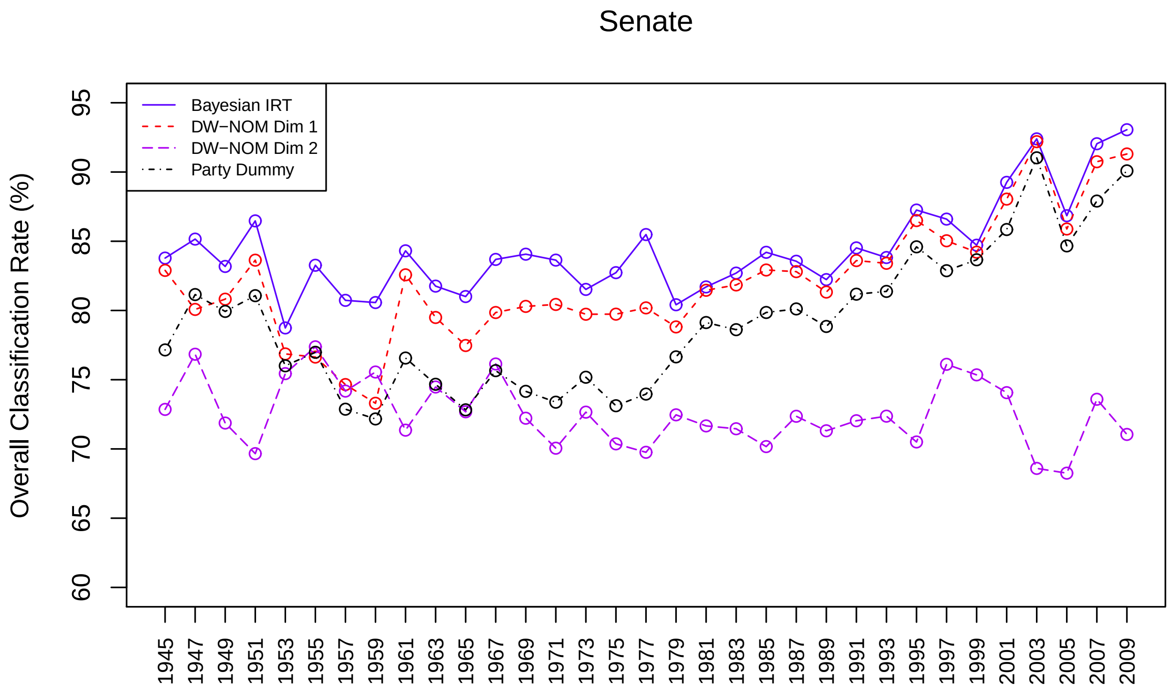



Figure 4: **Monte Carlo Analysis of IRT models and Interest Group Ratings.** The panels in the first column report results of weak polarization scenario. The second column reports results of strong polarization scenario. The majority size of votes selected by ADA varies across rows. In the second row, ADA ratings are generated using simulated votes that have a majority size of 60% or less; and the third rows, 80% or less. In all the cases, simulations were repeated 50 times.

**Table 1. Number and Percentage of Votes with Zero Discrimination Parameter.**

| | House | | | Senate | | |
|---|---|---|---|---|---|---|
| Congress | No of all foreign policy votes | No of votes with zero discrimimation parameter | % of votes with zero discrimination parameter | No of all foreign policy votes | No of votes with zero discrimimation parameter | % of votes with zero discrimination parameter |
| 79 | 52 | 2 | 3.85 | 61 | 3 | 4.92 |
| 80 | 25 | 0 | 0.00 | 48 | 2 | 4.17 |
| 81 | 42 | 2 | 4.76 | 137 | 5 | 3.65 |
| 82 | 60 | 1 | 1.67 | 81 | 2 | 2.47 |
| 83 | 27 | 1 | 3.70 | 61 | 7 | 11.48 |
| 84 | 20 | 0 | 0.00 | 63 | 14 | 22.22 |
| 85 | 33 | 5 | 15.15 | 49 | 6 | 12.24 |
| 86 | 26 | 2 | 7.69 | 78 | 11 | 14.10 |
| 87 | 36 | 0 | 0.00 | 86 | 9 | 10.47 |
| 88 | 40 | 0 | 0.00 | 108 | 12 | 11.11 |
| 89 | 38 | 0 | 0.00 | 86 | 12 | 13.95 |
| 90 | 62 | 1 | 1.61 | 101 | 3 | 2.97 |
| 91 | 71 | 1 | 1.41 | 121 | 10 | 8.26 |
| 92 | 106 | 2 | 1.89 | 243 | 23 | 9.47 |
| 93 | 161 | 5 | 3.11 | 193 | 22 | 11.40 |
| 94 | 183 | 6 | 3.28 | 136 | 12 | 8.82 |
| 95 | 243 | 7 | 2.88 | 209 | 14 | 6.70 |
| 96 | 234 | 8 | 3.42 | 165 | 16 | 9.70 |
| 97 | 128 | 5 | 3.91 | 142 | 8 | 5.63 |
| 98 | 164 | 8 | 4.88 | 175 | 15 | 8.57 |
| 99 | 113 | 2 | 1.77 | 194 | 11 | 5.67 |
| 100 | 208 | 5 | 2.40 | 207 | 14 | 6.76 |
| 101 | 150 | 7 | 4.67 | 130 | 16 | 12.31 |
| 102 | 131 | 1 | 0.76 | 112 | 8 | 7.14 |
| 103 | 131 | 4 | 3.05 | 118 | 11 | 9.32 |
| 104 | 188 | 8 | 4.26 | 136 | 6 | 4.41 |
| 105 | 167 | 5 | 2.99 | 84 | 12 | 14.29 |
| 106 | 125 | 1 | 0.80 | 101 | 15 | 14.85 |
| 107 | 123 | 2 | 1.63 | 107 | 7 | 6.54 |
| 108 | 188 | 4 | 2.13 | 136 | 7 | 5.15 |
| 109 | 219 | 2 | 0.91 | 94 | 6 | 6.38 |
| 110 | 217 | 1 | 0.46 | 71 | 2 | 2.82 |
| 111 | 121 | 0 | 0.00 | 49 | 3 | 6.12 |



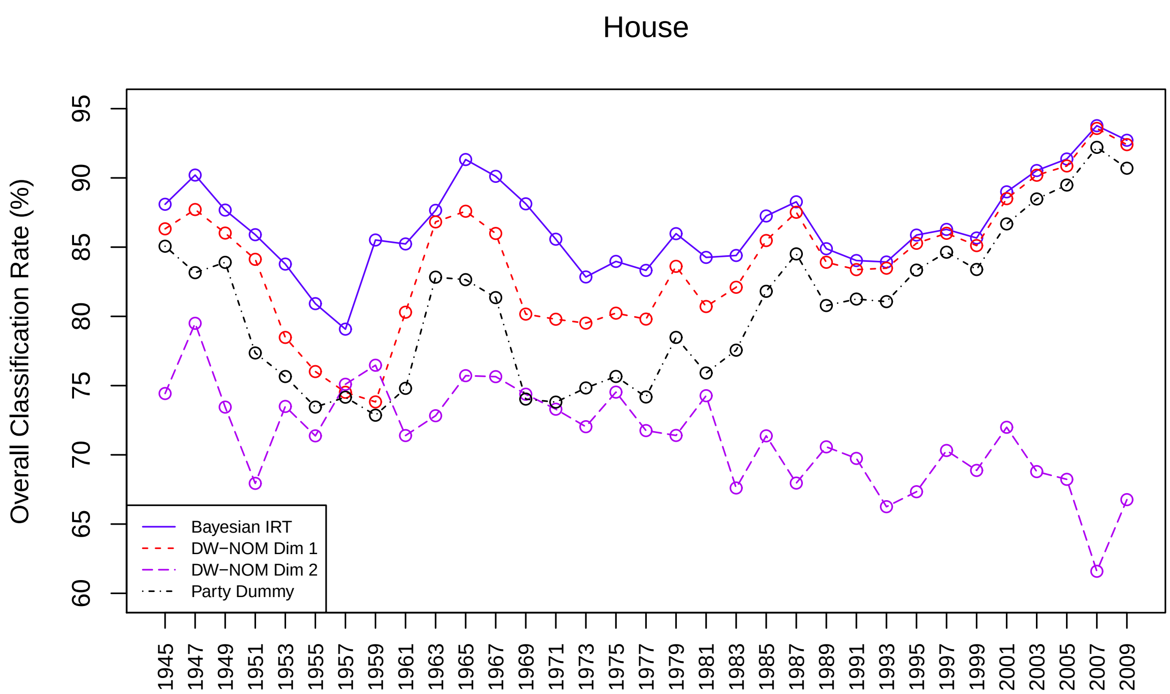




Figure 5: **Overall Classification Rates of Logit Models Fit to Foreign Policy Votes Using Each Measure as the Sole Independent Variable.** The classification rates represent the overall classification rates of logit models fit to each of the foreign policy votes in each Congress. Each logit model includes each one of the measures—the Bayesian IRT measure, the dimension 1 score of DW-NOMINATE, the dimension 2 score of DW-NOMINATE, and a party indicator variable (GOP=1)—as the sole predictor.

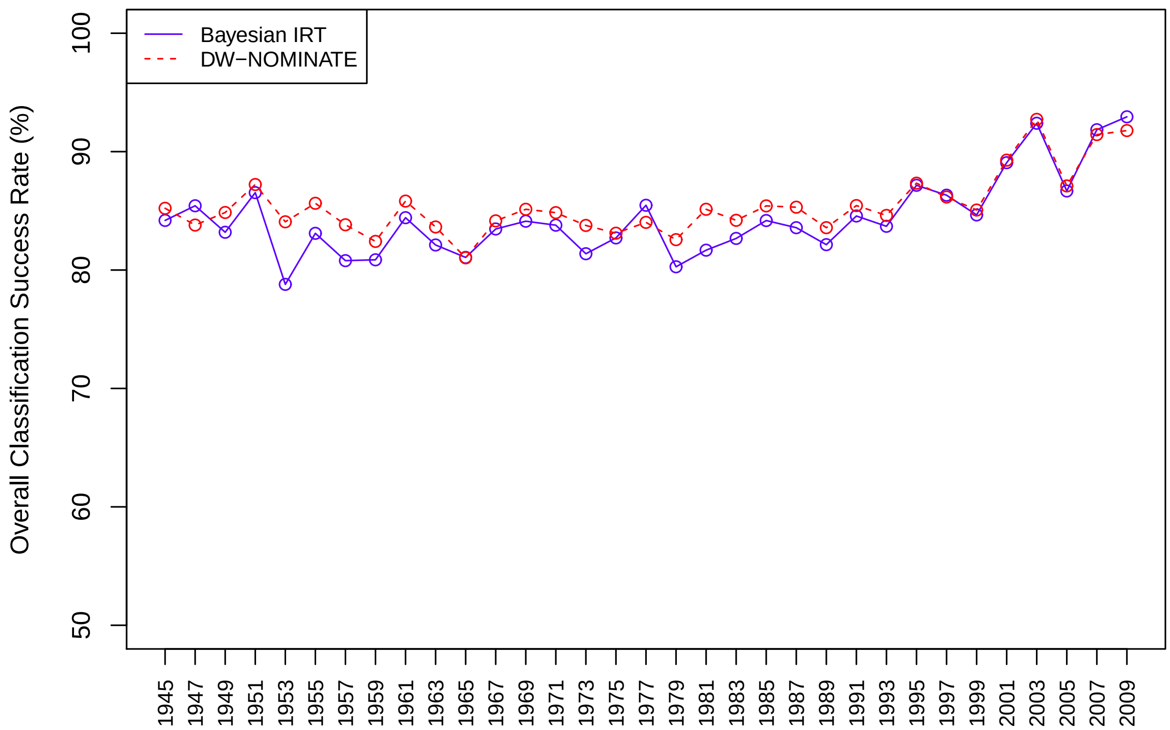
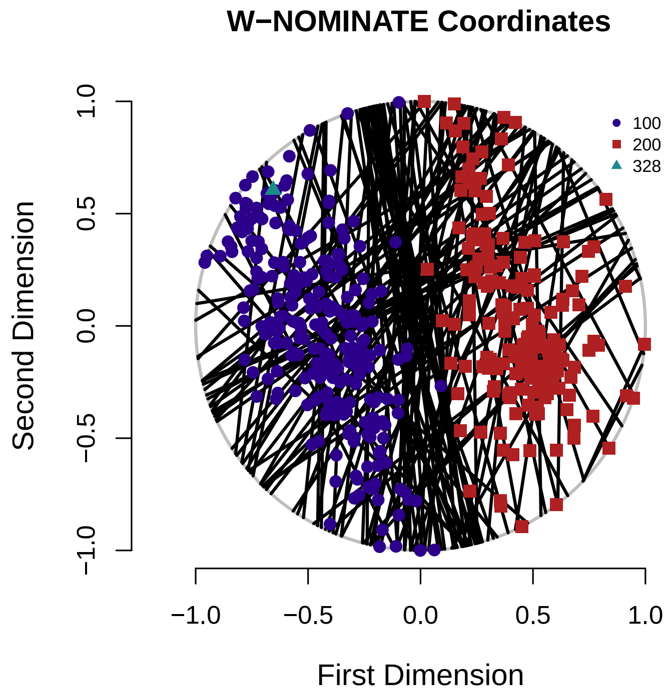



Figure 6: **Overall Classification Rates of the IRT and DW-NOMINATE Models on Foreign Policy Votes in the Senate.** The classification rates for DW-NOMINATE are obtained from *vote-view.com*.



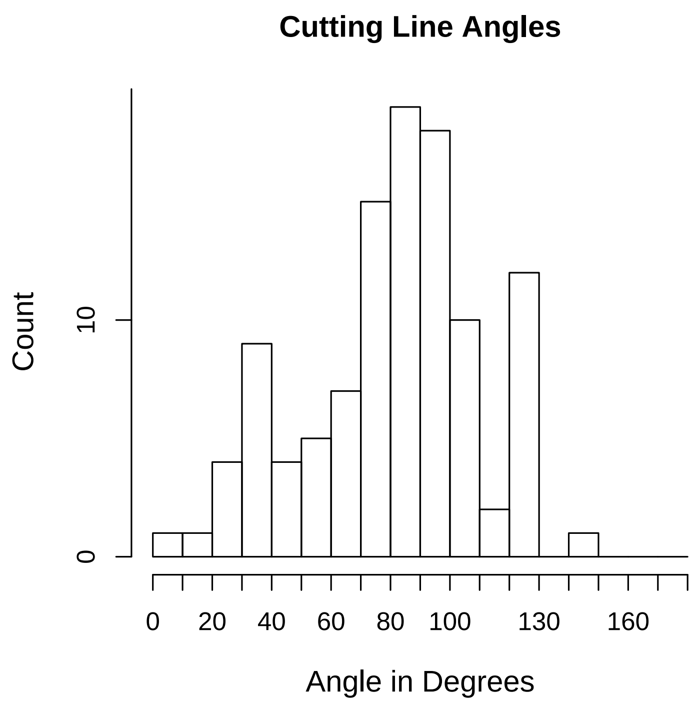




Figure 7: **103rd House W-NOMINATE Scores with Cutlines (Top) and Cutline Angles (Bottom) for all Foreign Policy Votes.** The top panel plots W-NOMINATE scores with cutlines for all foreign policy votes. The bottom panel shows the distribution of cutline angles for foreign policy votes. 17% (18 out of 108) of the foreign policy votes have cutline angles below 45 or above 135 degree, indicating that these votes split members more along the second dimension than the first dimension.

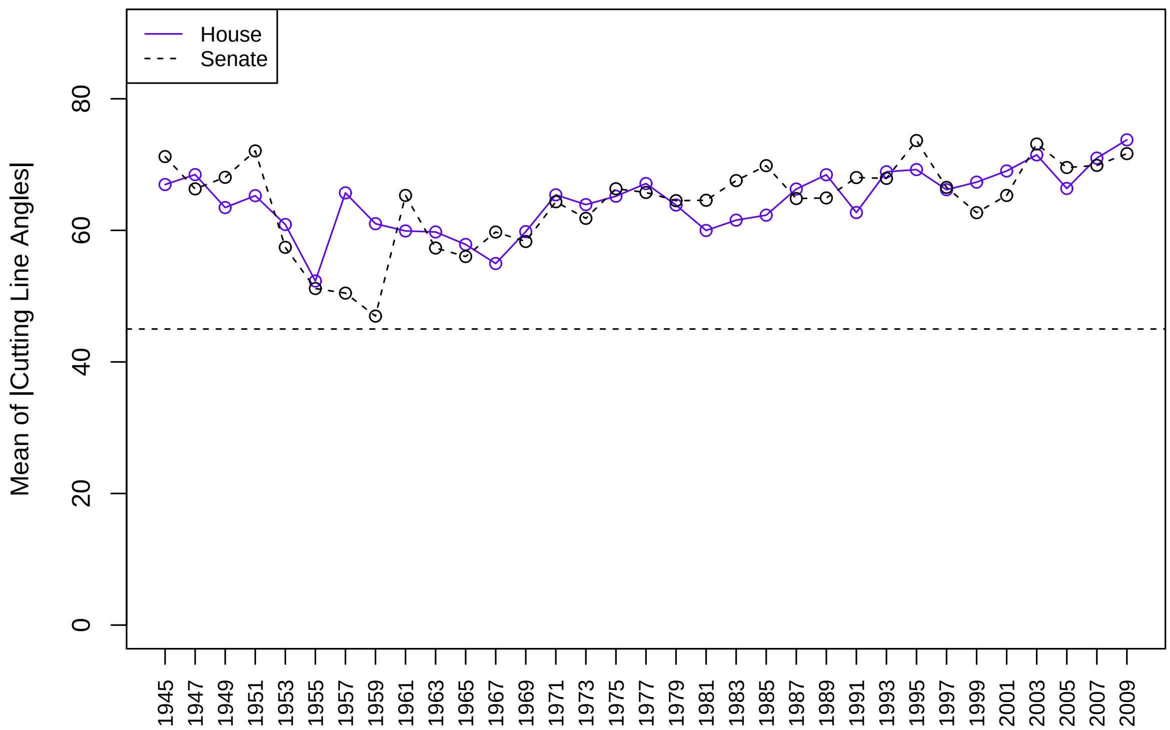



Figure 8: **Average Cutline Angle Degrees for DW-NOMINATE Roll Call Output of Foreign Policy Votes.** Each dot represents the average of the absolute values of cutline angle degrees of DW-NOMINATE measure of foreign policy votes in each Congress. The dotted line represents 45 degrees. In this Voteview file, cutting line angle degrees range from 90 to -90 degrees. Therefore, an absolute value above 45 degrees indicates that the vote splits more members along dimension 1 than dimension 2. Data Source: voteview.com.

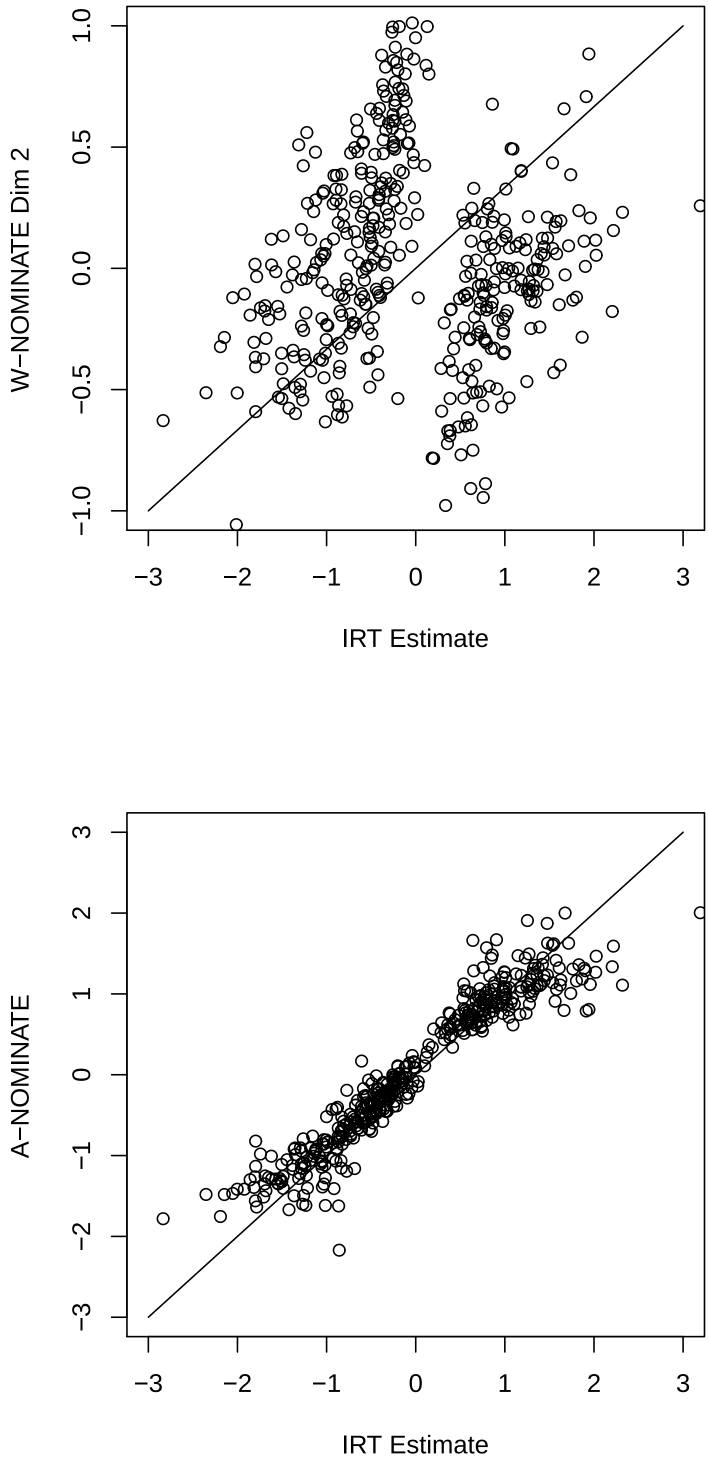



Figure 9: **Comparison of IRT estimate, the dimension 2 score of W-NOMINATE, and A-NOMINATE.** The location of each dot on the horizontal dimension represents the IRT estimate of a legislator. The location of the vertical dimension represents the legislator's W-NOMINATE dimension 2 score and A-NOMINATE score.

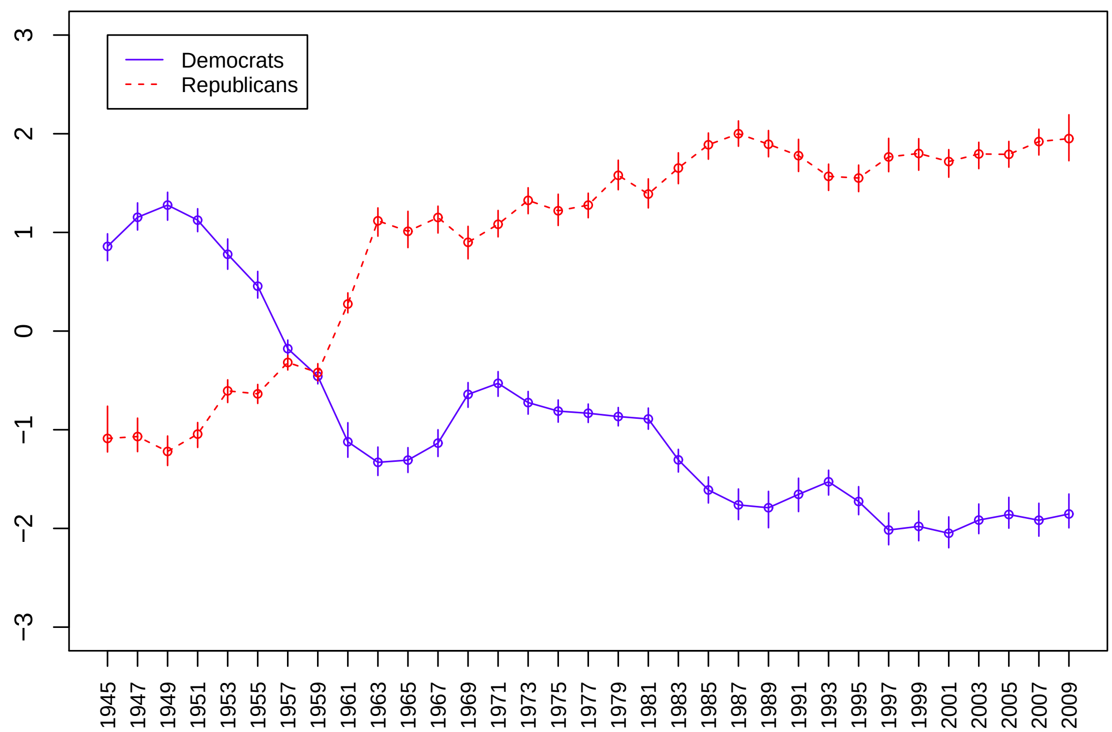



Figure 10: **House Party Medians.** The data are all foreign and defense policy votes from the House since 1945. Each party's median represents the median ideal point of each party's members.The bars around the dots represent the 95% credible intervals.

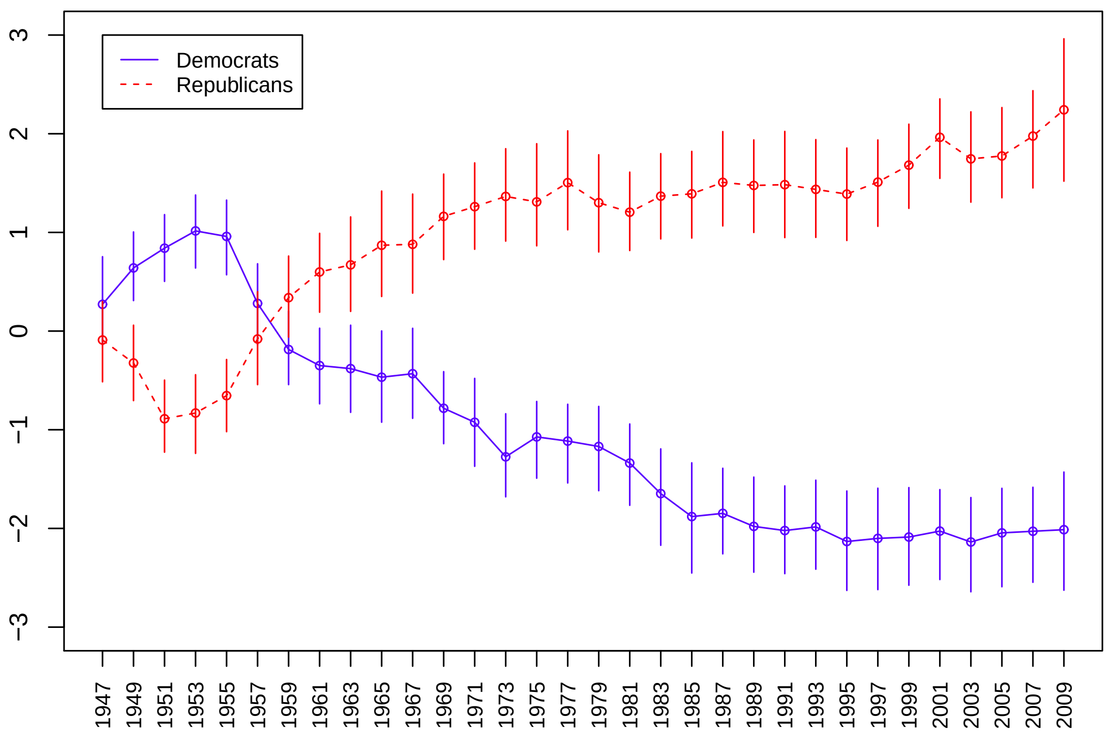



Figure 11: **Senate Party Medians (CQ Key Votes).** The data are foreign and defense policy Key Votes identified by *Congressional Quarterly*. Each party's median represents the median ideal point of each party's members.The bars around the dots represent the 95% credible intervals.

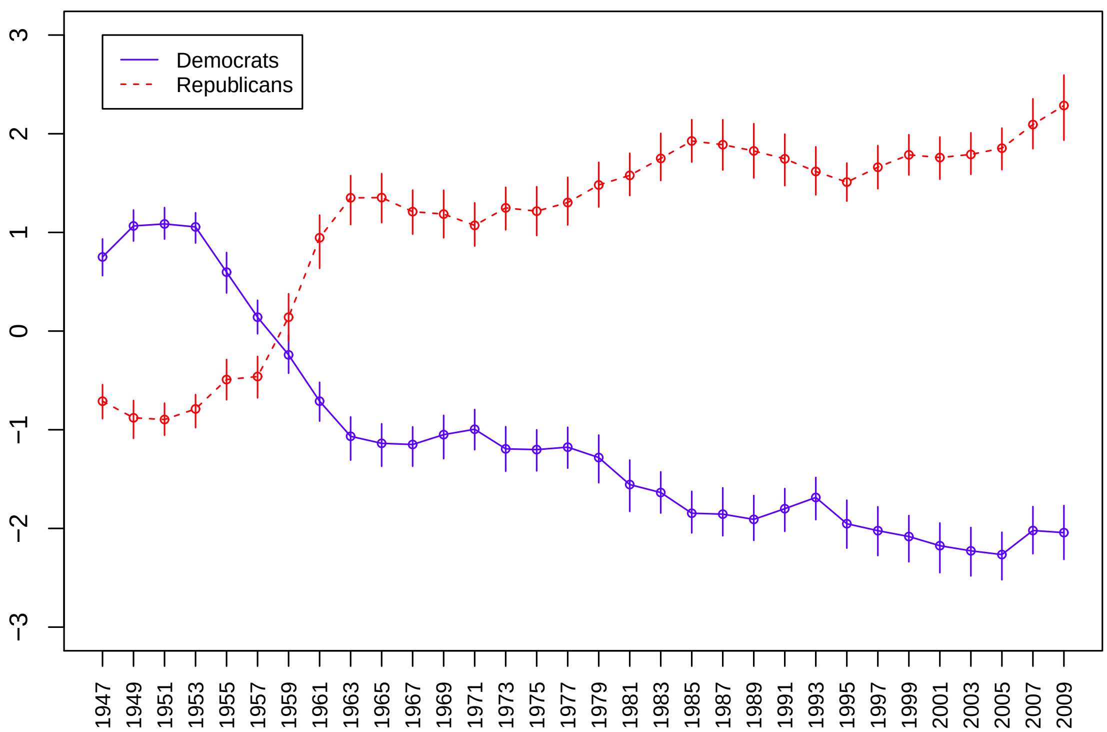



Figure 12: **House Party Medians (CQ Key Votes).** The data are foreign and defense policy Key Votes identified by *Congressional Quarterly*. Each party's median represents the median ideal point of each party's members.The bars around the dots represent the 95% credible intervals.

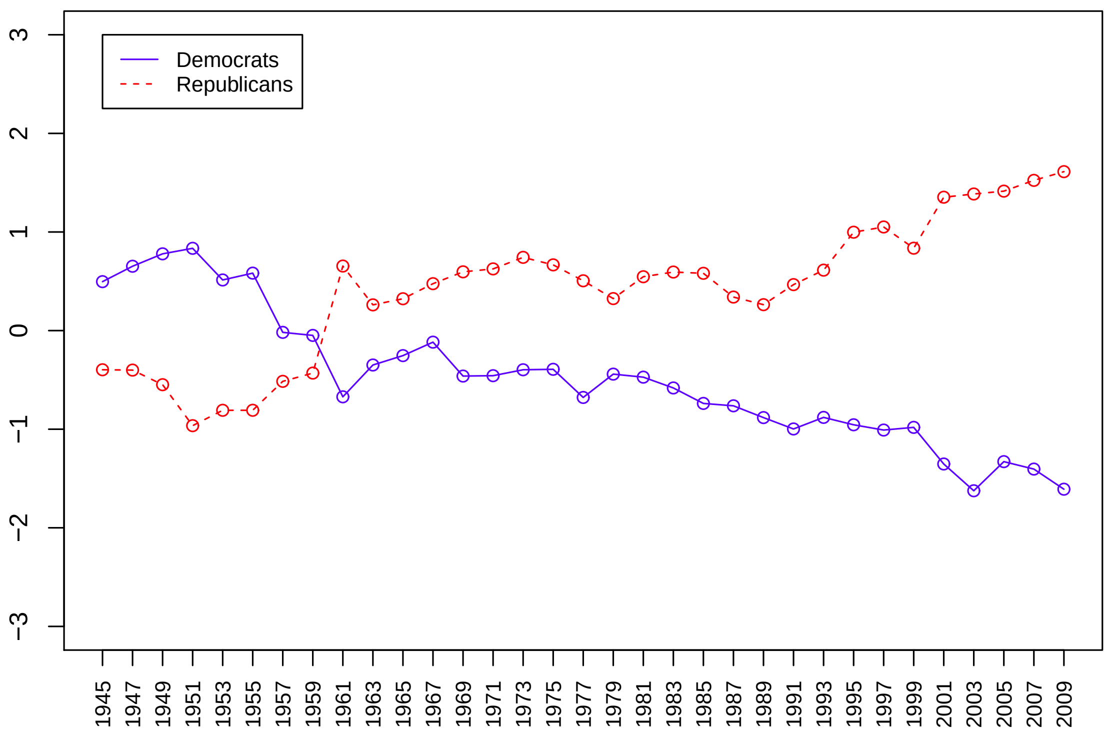



Figure 13: **Senate Party Medians (Trade and Immigration Votes Excluded).** The data are all foreign and defense policy votes, except for all trade and immigrations votes, from the Senate since 1945. Each party's median represents the median ideal point of each party's members.The bars around the dots represent the 95% credible intervals.

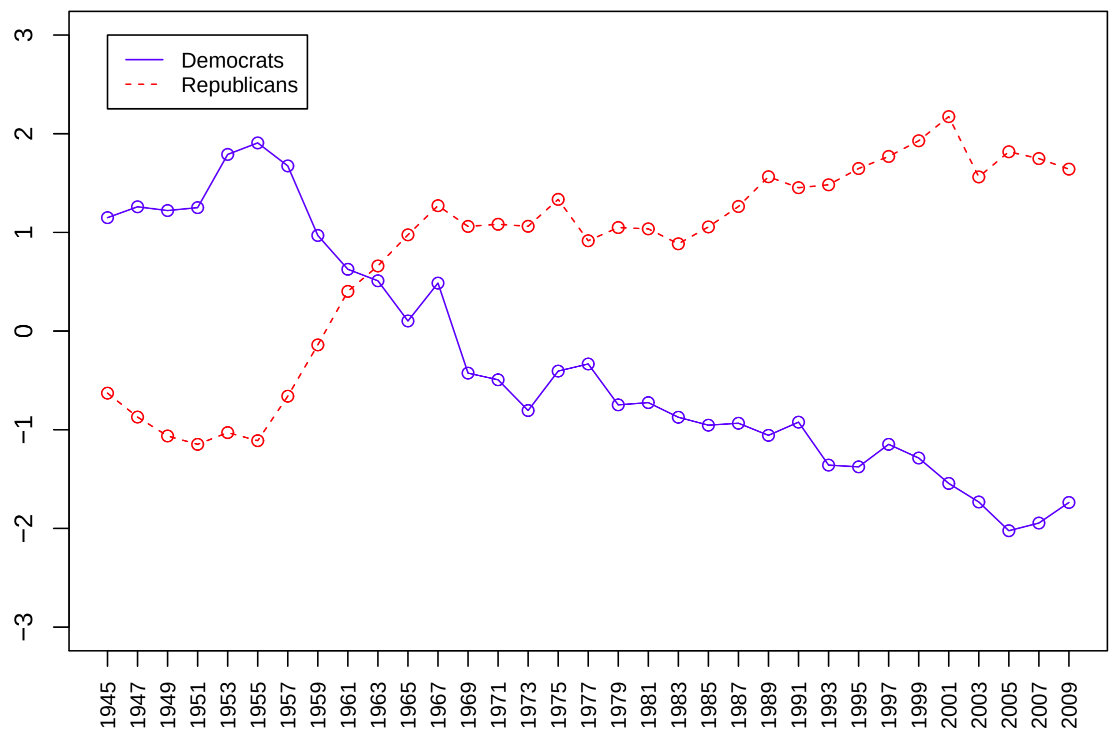



Figure 14: **Senate Party Medians (Defense Spending Votes Only).** The data are defense spending votes only from the Senate since 1945. Each party's median represents the median ideal point of each party's members.The bars around the dots represent the 95% credible intervals.

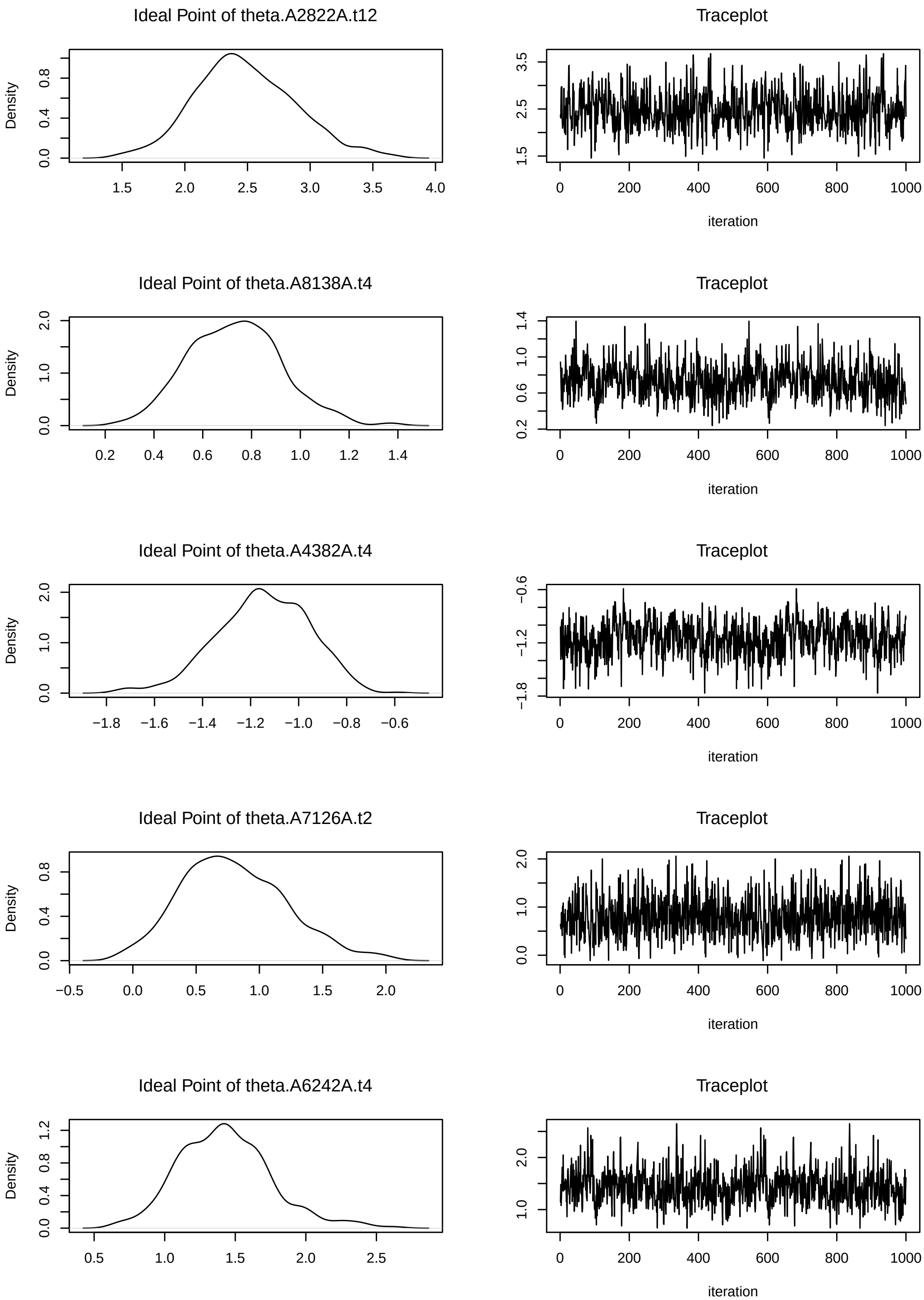



Figure 15: **Density and Trace Plots for a Sample of Ideal Point Estimates.** The MCMC chains are for ideal points of Senators in dynamic IRT estimation.

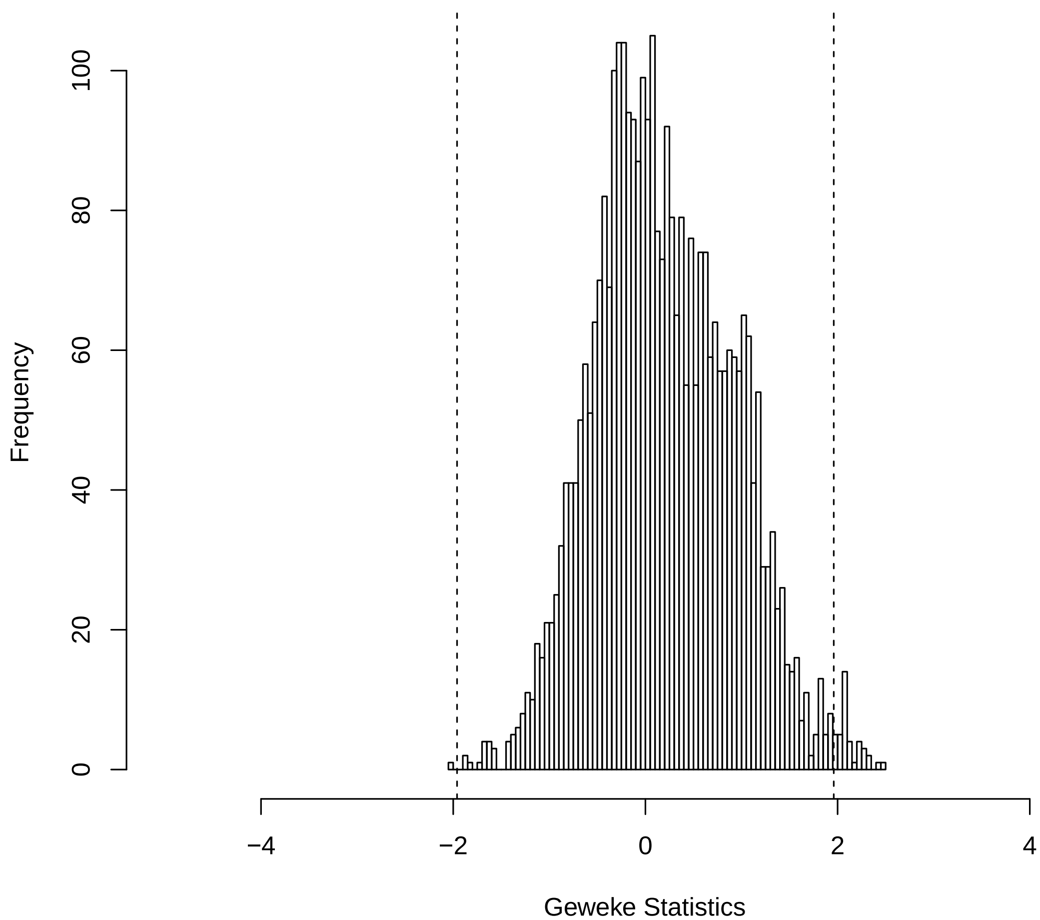



Figure 16: **Histogram of the Geweke Statistics.** The dotted lines represent -1.96 and 1.96. 1.2% of the statistics are outside the range.