Appendix for Absence Makes the Ideal Points Sharper: Full-data IRT Models for Legislatures

Robert Kubinec

Department of Politics

University of Virginia

September 29, 2017

Complete Exposition of Absence-Inflated IRT

To write out the model fully, we need to address the four possible outcomes y_{ijkr} for each legislator i and bill j: $\{N, A_b, Y, A_s\}$ where Y stands for yes votes, A_b for abstentions, N for no votes, and A_s for absences. I denote the probability that a legislator is absent P_r. We can the probability $f(y_{ijkr})$ of each the four outcomes as functions of P_r and the probabilities of two of the three possible votes, P_N, P_{Ab}, where the third vote P_Y is the residual probability $1 - P_N - P_{Ab}$.

$$f(y_{ijkr}) = \begin{cases} P_r \text{ if } y_{ijkr} = A_s, \\ (1 - P_r)P_N \text{ if } y_{ijkr} = N, \\ (1 - P_r)P_{Ab} \text{ if } y_{ijkr} = A_b \\ (1 - P_r)(1 - P_N - P_{Ab}) \text{ if } y_{ijkr} = Y \end{cases}$$

This version of the model shows clearly how the probability of absence P_r deflates the probabilities of each of the vote outcomes equally. As the probability of absence increases, the probability of each vote outcome falls proportionally. We can combine these expressions into a single likelihood by denoting the outcomes by an indicator function I subscripted with the category and multiplying over the legislators and votes:

$$L(P_r, P_N, P_{Ab}, P_Y|y_{ijkr}) = \prod_{i=1}^{I}\prod_{j=1}^{J} P_r^{I_{As}}\big((1-P_r)P_N\big)^{I_N}\big((1-P_r)P_{Ab}\big)^{I_{Ab}}\big((1-P_r)(1-P_N-P_{Ab}\big)^{I_Y}$$

We then substitute the probabilities P_r, P_N, P_{Ab}, and P_Y with logistic link functions to incorporate the IRT model for each specific category to show the summation over i and j. To do so we also incorporate cutpoints c_1 and c_2 that represent the ordered structure of the votes N, A_b and Y.

$$\begin{aligned}
P_r &= \frac{1}{1+e^{x_i'\gamma_j-\omega_j}} \\
P_N &= 1 - \frac{1}{1+e^{x_i'\alpha_j-\beta_j}} \\
P_{Ab} &= \Big(\frac{1}{1+e^{x_i'\alpha_j-\beta_j}} - c_2\Big) - \Big(\frac{1}{1+e^{x_i'\alpha_j-\beta_j}} - c_1\Big) \\
P_Y &= \frac{1}{1+e^{x_i'\alpha_j-\beta_j}} - c_2
\end{aligned}$$

This formulation makes it very clear how the cutpoints affect the model. The abstention category is specifically the difference between the two cutpoints, while the yes category is all the probability mass above the higher cutpoint c_2. The no category is the remainder after the probabilities of abstention and yes votes have been calculated.