

ABSENCE MAKES THE IDEAL POINTS SHARPER: FULL-DATA IRT MODELS FOR LEGISLATURES

Robert Kubinec

Woodrow Wilson Department of Politics

Existing IRT 2-PL for Rollcall Data

Existing ideal point models based on IRT (Clinton, Jackman, & Rivers, 2004) use the following likelihood:

$$L(\beta, \alpha, X|Y) = \prod_n \prod_m \prod_{i=1} \prod_{j=1} \zeta(x'_i \beta_j - \alpha_j)^{y_{ij}} \times (1 - \zeta(x'_i \beta_j - \alpha_j))^{(1-y_{ij})}$$

Where β_j represent bill discrimination parameters, α_j represent bill intercepts, x_i are the legislator ideal points and ζ is a link function (generally logit or probit).

- Model assumes that legislators are present at all votes.
- Model assumes that all votes are binary (yes or no).

Missing Data?

This basic IRT model has been criticized for failing to meet core assumptions about how legislators vote.

- Imputation of a missing latent variable is difficult conceptually and practically. In particular, all the assumptions of imputation, such as conditionally missing at random (CMAR), should be met (Little & Rubin, 2002).
- I argue that all legislator actions can be interpreted as strategic behavior and can be included without imputation through an appropriate likelihood.

Absence-Inflated Ordinal IRT

I propose a likelihood that implements a distinct outcome for the $K = 3$ possible categories of votes (no, abstain, yes) and the $R \in \{0, 1\}$ binary decision to be absent or present. Adopting notation from Stan Development Team (2016), I use an ordered logistic likelihood for vote choice when a legislator is present ($r = 1$) such that each outcome is separated by $c \in \mathbb{R}^{K-1}$ cutpoints:

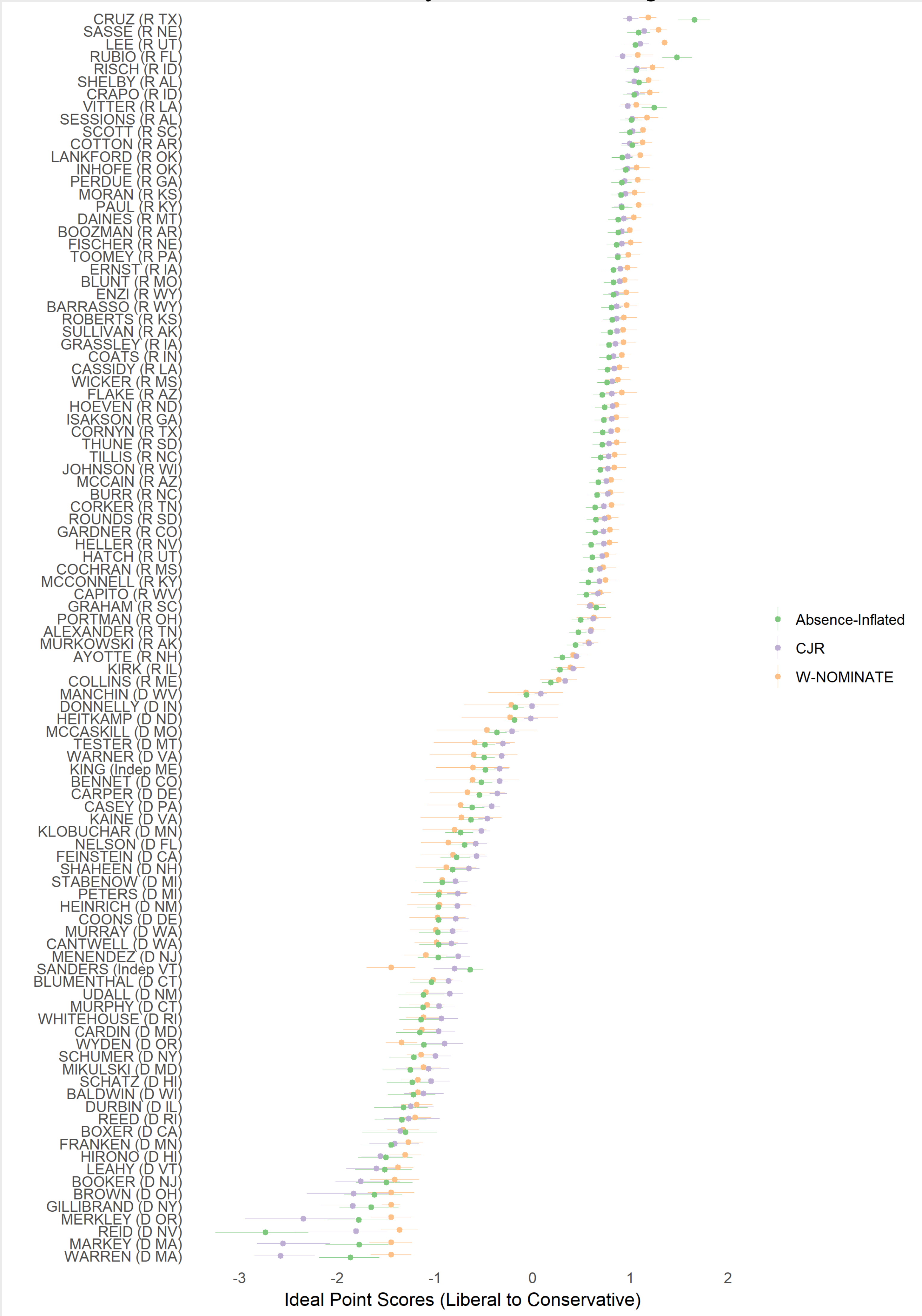
$$L(\beta, \alpha, X|Y_k) = \prod_n \prod_m \prod_{i=1} \prod_{j=1} \begin{cases} 1 - \zeta(x'_i \beta_j - \alpha_j - c_1) & \text{if } K = 0 \\ \zeta(x'_i \beta_j - \alpha_j - c_{k-1}) - \zeta(x'_i \beta_j - \alpha_j - c_k) & \text{if } 0 < k < K, \text{ and} \\ \zeta(x'_i \beta_j - \alpha_j - c_{k-1}) - 0 & \text{if } k = K \end{cases}$$

This ordinal model is deflated by the probability that a legislator is present to model the combined $Y_{(k,r)}$ likelihood. The probability of $r = 0$ absence is a separate IRT model of common ideal points x_i , and separate bill discrimination/intercepts γ_j and ω_j .

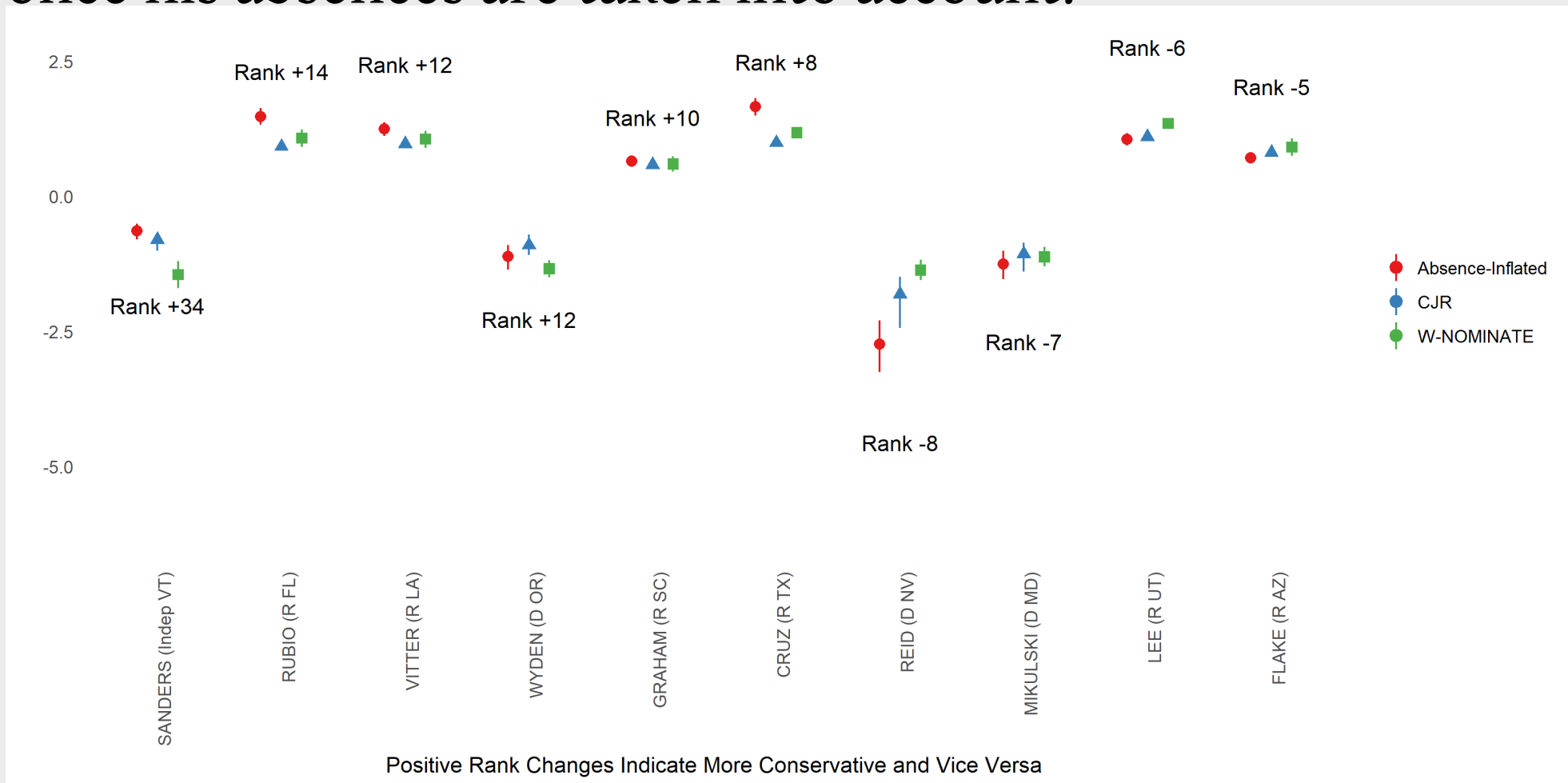
$$L(\beta, \alpha, X, Q, \gamma, \omega|Y_{kr}) = \prod_n \prod_m \prod_{i=1} \prod_{j=1} \begin{cases} \zeta(x'_i \gamma_j - \omega_j) & \text{if } r = 0, \text{ and} \\ (1 - \zeta(x'_i \gamma_j - \omega_j)) L(\beta, \alpha, X|Y_{k1}) & \text{if } r = 1 \end{cases}$$

Case Study: 114th Senate

In the most recent Senate, modeling absences as an explicit category still largely aligns with the NOMINATE and `ideal` models that are popularly used for ideal point estimation. This is because of low absence rates (and virtually no absentions) among U.S. Senators.

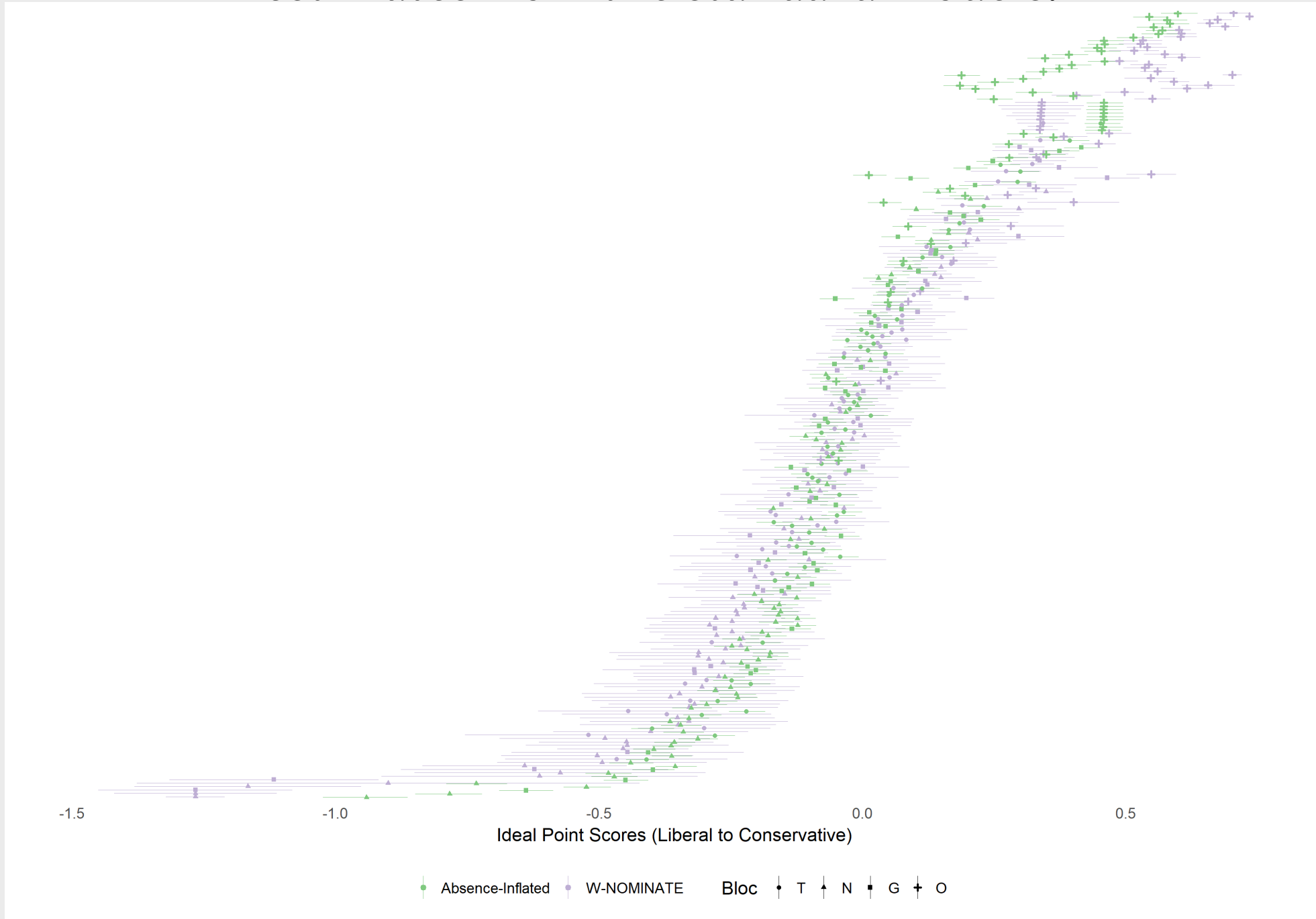


However, ideal points do tend to differ for Senators who were on the campaign trail and had to make strategic decisions about which bills to show up for. Ted Cruz and Marco Rubio all appear far more conservative once their absences are taken into account, while Bernie Sanders is somewhat more moderate. On the other hand, Harry Reid is far more liberal once his absences are taken into account.



Case Study: Tunisian Parliament

Parliaments with coalition governments can have much higher absentee rates and abstention rates. The data below from Tunisia's legislative assembly, which is considered a transitional democracy, show significantly diverging estimates from the standard models.



R Package `idealstan`

This model and associated plots, along with other IRT ideal point models, are available via the Github-based R package `idealstan` (www.github.com/saudiwin/idealstan). This package interfaces with the Stan MCMC library for full and approximate Bayesian inference.