A Novel Class of Unfolding Models for Binary Preference Data

Rayleigh Lei[1]   and Abel Rodríguez[2]

[1]Institute for Social Research, University of Michigan, Ann Arbor, MI, USA; [2]Department of Statistics, University of Washington, Seattle, WA, USA

Corresponding author: Rayleigh Lei; Email: rlei13@uw.edu

Abstract

We develop a new class of spatial voting models for binary preference data that can accommodate both monotonic and non-monotonic response functions, and are more flexible than alternative "unfolding" models previously introduced in the literature. We then use these models to estimate revealed preferences for legislators in the U.S. House of Representatives and justices on the U.S. Supreme Court. The results from these applications indicate that the new models provide superior complexity-adjusted performance to various alternatives and that the additional flexibility leads to preferences' estimates that more closely match the perceived ideological positions of legislators and justices.

1. Introduction

Methods for estimating the preferences of members of deliberative bodies from their voting records have become a foundational tool in political science. Techniques, such as NOMINATE (Poole and Rosenthal 1985), IDEAL (Jackman 2001), and its relatives, are used for exploratory purposes to describe the behavior of voters (see, e.g., Jenkins 2006; Luque and Sosa 2023; Poole and Rosenthal 2006; Shor and McCarty 2011), as well as in confirmatory settings to test specific theories of legislative or judicial behaviour (e.g., Clark 2012; Schickler 2000; Schwindt-Bayer and Corbetta 2004). However, despite their widespread use, these scaling methods have a number of limitations (see, e.g., Clinton 2012; Hug 2010; Jessee and Theriault 2014; Roberts 2007). In particular, in situations where the policy space is assumed to be unidimensional and it is common for voters on both ends of the political spectrum to vote together against those in the middle, standard methods often lead to what could be considered inaccurate estimates for the most extreme legislators (see, e.g., Lewis 2019a, 2019b; Yu and Rodriguez 2021). In response to these issues, Duck-Mayr and Montgomery (2023) recently introduced a Bayesian version of the generalized graded unfolding model (GGUM; Roberts, Donoghue, and Laughlin 2000) to the political science literature, and demonstrated that this Bayesian GGUM (BGGUM) is more appropriate than traditional scaling methods for ends-against-middle votes.

GGUMs originated in the psychology literature, where they have arguably become the most popular ideal point model for non-cognitive measurements that indicate how well an item describes the respondent's typical preference on a latent continuum. The key insight behind their construction is that, in one-dimensional settings, individuals might disagree with a particular statement when their own preferences lie too far in either direction from a reference point, leading to non-monotonic response functions. Roberts, Donoghue, and Laughlin (2000) describe this phenomenon as individuals potentially disagreeing either "from above" and "from below," and construct their model for the observed preferences by "folding" a traditional latent, "subjective" scale with a monotonic response function.

In our view, one shortcoming of the Roberts, Donoghue, and Laughlin (2000) and Duck-Mayr and Montgomery (2023) approach in the context of political science applications is the lack of a clear, explicit connection with spatial voting models (Davis, Hinich, and Ordeshook 1970; Enelow and Hinich 1984). A second challenge with the BGGUM is computational. While various packages implementing algorithms to fit GGUMs exist (see, e.g., Tendeiro and Castro-Alvarez 2019; Tu *et al.* 2021; Wang, de la Torre, and Drasgow 2015), estimation can be challenging in practice. In particular, in the context of Bayesian inference, many of the methods are based on Metropolis–Hastings algorithms that rely on proposals that need to be carefully calibrated and can fail to properly explore the posterior distribution. Duck-Mayr and Montgomery (2023) present a Metropolis-coupled MC3 algorithm that uses a set of parallel chains running at various temperatures. This approach shows improved mixing when compared with previous algorithms. However, their algorithm still relies on a series of random-walk Metropolis–Hastings proposals that need to be carefully calibrated for each new dataset, and our numerical experiments suggest that these calibration can be difficult and time-consuming. Furthermore, the structure of the MC3 algorithm makes extensions of the basic model to hierarchical settings difficult to implement, especially if they involve model comparisons or other discrete random variables (see, e.g., Lofland, Rodriguez, and Moser 2017; Moser, Rodríguez, and Lofland 2021; Rodriguez and Moser 2015). One final challenge refers to prior specification for BGGUMs. The most common choice of priors, introduced in De La Torre, Stark, and Chernyshenko (2006) and adopted by Duck-Mayr and Montgomery (2023), involves the use of four-parameter Beta distributions with a compact support. There is no clear guidance in the literature on how the hyperparameters of these priors should be chosen beyond vague warnings about retrospectively ensuring that the support of the priors contains the support of the corresponding posteriors.

In this paper, we introduce an alternative to the GGUM of Roberts, Donoghue, and Laughlin (2000) whose construction relies on the random utility framework of McFadden (1973), providing an explicit link between unfolding models and spatial voting models. We also introduce an efficient Gibbs sampling algorithm that does not require ad hoc tuning, and discuss criteria for prior specification. The resulting model is used to analyze voting data from the U.S. House of Representatives between 1987 and 2022. In this context, we show that our new unfolding model provides a better complexity-adjusted fit to the data than IDEAL and the BGGUM of Duck-Mayr and Montgomery (2023) (as measured by the Watanabe–Akaike Information Criteria [WAIC]; see e.g., Watanabe 2013; Watanabe and Opper 2010), as well as estimates of legislators preferences that better match what would seem to be their true ideological leanings. We then move to extend this basic framework to create *dynamic* unfolding models in which the latent traits are allowed to evolve over time. The resulting model, which can be seen as a generalization of the dynamic factor model for binary data introduced by Martin and Quinn (2002), is used to analyze voting patterns in U.S. Supreme Court between 1935 and 2021. In this context, we show that voting patterns at key points of the U.S. Supreme Court can be better described by our new dynamic unfolding model.

2. A Spatial Formulation for Unfolding Models

Let $y_{i,j}$ represent the vote of member $i = 1, 2, \ldots, I$ on issue $j = 1, 2, \ldots, J$. We assume that there is a latent one-dimensional Euclidean space (the *policy space*) and that each voter has a preferred position in that space (their *ideal point*), denoted by $\beta_1, \ldots, \beta_I$. Additionally, each vote has associated with it three positions, $\psi_{j,1}$, $\psi_{j,2}$, and $\psi_{j,3}$, such that $\psi_{j,2}$ corresponds to the preferred position for a positive ("Aye") vote on issue j, while $\psi_{j,1}$ and $\psi_{j,3}$ are the preferred positions for a negative ("Nay") vote (please see Figure 1). In the sequel, we assume that either $\psi_{j,1} < \psi_{j,2} < \psi_{j,3}$ (in which case $\psi_{j,1}$ and $\psi_{j,3}$ correspond to a negative vote because the member disagrees "from below" and "from above" respectively) or $\psi_{j,3} < \psi_{j,2} < \psi_{j,1}$ (in which $\psi_{j,1}$ and $\psi_{j,3}$ have the opposite interpretation). Under the order constraint, we can interpret one of the "Nay" positions (either $\psi_{j,1}$ or $\psi_{j,3}$, depending on the orientation of the latent scale) as the *status quo*, the "Aye" position $\psi_{j,2}$ as the proposed policy, and the second "Nay" position as an even more extreme policy that some legislators might find preferable to the proposed one. This interpretation of the latent positions aligns with a common narrative behind
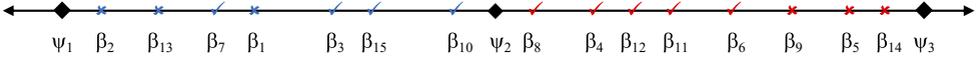
Figure 1. Cartoon representation of the spatial voting model construction of our unfolding model. Squares represent the ψ_j's, check marks represent the ideal points of legislators voting in favor the issue, and crosses represent the ideal points of legislators voting against it.

extreme-against-the-middle votes in recent U.S. Congresses in which the most extreme members of the majority party vote against their own party's proposed policy because it is not extreme enough and they are not willing to settle for incremental change.

Similar to Jackman (2001), we assume that members choose among these three options on the basis of quadratic utility functions that depend on the distance between their ideal point β_i and the vote positions $\psi_{j,1}$, $\psi_{j,2}$ and $\psi_{j,3}$,

$$U_{N^-}(\beta_i, \psi_{j,1}) = -(\beta_i - \psi_{j,1})^2 + \epsilon_{i,j,1},$$
$$U_Y(\beta_i, \psi_{j,2}) = -(\beta_i - \psi_{j,2})^2 + \epsilon_{i,j,2},$$
$$U_{N^+}(\beta_i, \psi_{j,3}) = -(\beta_i - \psi_{j,3})^2 + \epsilon_{i,j,3},$$

where $\epsilon_{i,j,t,1}$, $\epsilon_{i,j,t,2}$, and $\epsilon_{i,j,3}$ are independent and identically distributed shocks. Then, an affirmative vote occurs if and only if $U_Y > U_{N^-}$ and $U_Y > U_{N^+}$, so that

$$P(y_{i,j} = 1 \mid \beta_i, \psi_{j,1}, \psi_{j,2}, \psi_{j,3}) = P(U_Y(\beta_i, \psi_{j,2}) > U_{N^-}(\beta_i, \psi_{j,1}), U_Y(\beta_i, \psi_{j,2}) > U_{N^+}(\beta_i, \psi_{j,3}))$$
$$= P(\epsilon_{i,j,1} - \epsilon_{i,j,2} < \alpha_{j,1}(\beta_i - \delta_{j,1}), \epsilon_{i,j,3} - \epsilon_{i,j,2} < \alpha_{j,2}(\beta_i - \delta_{j,2})),$$

where $\alpha_{j,1} = 2(\psi_{j,2} - \psi_{j,1})$, $\alpha_{j,2} = 2(\psi_{j,2} - \psi_{j,3})$, $\delta_{j,1} = (\psi_{j,1} + \psi_{j,2})/2$, and $\delta_{j,2} = (\psi_{j,3} + \psi_{j,2})/2$.

In the special case where $\epsilon_{i,j,k}$'s are independently and identically distributed from a standard Gumbel distribution, standard theory indicates that

$$P(y_{i,j} = 1 \mid \beta_i, \alpha_{j,1}, \delta_{j,1}, \alpha_{j,2}, \delta_{j,2}) = \frac{1}{1 + \exp\{-\alpha_{j,1}(\beta_i - \delta_{j,1})\} + \exp\{-\alpha_{j,2}(\beta_i - \delta_{j,2})\}}. \tag{1}$$

Equation (1) is very similar (although not identical) to that associated with the GGUM for binary data in Roberts, Donoghue, and Laughlin (2000) and Duck-Mayr and Montgomery (2023).

Alternatively, if the shocks are normally distributed, we obtain a probit unfolding model with response function

$$P(y_{i,j} = 1 \mid \beta_i, \alpha_{j,1}, \delta_{j,1}, \alpha_{j,2}, \delta_{j,2}) = \int_{-\infty}^{\alpha_{j,1}(\beta_i - \delta_{j,1})} \int_{-\infty}^{\alpha_{j,2}(\beta_i - \delta_{j,2})} \frac{1}{2\sqrt{3}\pi} \exp\left\{-\frac{1}{3}(z_1^2 - z_1 z_2 + z_2^2)\right\} \, dz_1 dz_2. \tag{2}$$

This is simply the cumulative distribution function of the bivariate normal distribution with mean $\mathbf{0}$, and covariance matrix $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ evaluated at $(\alpha_{j,1}(\beta_i - \delta_{j,1}), \alpha_{j,2}(\beta_i - \delta_{j,2}))$. In the remainder of this paper, we will focus on this probit version of the unfolding model.

An appealing property of our construction is that it clearly includes IDEAL as a special case. For example, under the assumption of Gaussian shocks, and if we have $\psi_{j,1} < \psi_{j,2} < \psi_{j,3}$, letting $\psi_{j,3} \to \infty$ implies that

$$P(y_{i,j} = 1 \mid \beta_i, \alpha_{j,1}, \delta_{j,1}, \alpha_{j,2}, \delta_{j,2}) \to \Phi(\alpha_{j,1}(\beta_i - \delta_{j,1})), \tag{3}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution.

2.1. Prior Distributions

We adopt a Bayesian approach to inference and proceed to discuss the prior distributions associated with the unknown model parameters $\beta_1, \ldots, \beta_I$, $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_J$ and $\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_J$. We first discuss the functional form

of the priors, which is chosen in part for computational convenience, and then discuss hyperparameter selection.

The choice for the prior distribution on the ideal points $\beta_1, \ldots, \beta_I$ is straightforward. We follow most of the literature and let the ideal points be independent and identically distributed from a standard normal distribution, $\beta_i \sim N(0,1)$. Fixing the mean and variance of this distribution helps in addressing some of the identifiability issues associated with spatial voting models (see Section 2.2).

The design of priors for the vote-specific parameters $\boldsymbol{\alpha}_j = (\alpha_{j,1}, \alpha_{j,1})'$ and $\boldsymbol{\delta}_j = (\delta_{j,1}, \delta_{j,2})'$ is more difficult. One consideration to keep in mind is that there is no natural "correct" direction for the underlying policy space, so it is important that $\alpha_{j,1}$ and $\alpha_{j,2}$ marginally be allowed to have full support over the whole real line. A second consideration is that our original construction requires either $\psi_{j,1} < \psi_{j,2} < \psi_{j,3}$ or $\psi_{j,3} < \psi_{j,2} < \psi_{j,1}$. In addition to facilitating the interpretation of the model, there are at least two more reasons why introducing this constraint is important. The first one is identifiability. Indeed, note that if $\psi_{j,1}$ and $\psi_{j,3}$ (the "Nay" positions) are on the "same side" of $\psi_{j,2}$ (the "Aye" position), there is no way to separately learn both $\psi_{j,1}$ and $\psi_{j,3}$ from the data since they are observationally equivalent. Connected to this, we note that the response function is automatically monotonic when $\psi_{j,1}$ and $\psi_{j,3}$ are on the same side of $\psi_{j,2}$. Hence, without the order constraint, there would be two ways to represent "traditional" partisan votes (which correspond to monotonic response functions): (a) by placing both $\psi_{j,1}$ and $\psi_{j,3}$ on the same side of $\psi_{j,2}$, or (b) by putting them on opposite sides of $\psi_{j,2}$ but making one of them very extreme (recall Equation (3)). This would lead to a multimodal posterior distribution, which would be challenging to explore. Hence, our priors force $\alpha_{j,1}$ and $\alpha_{j,2}$ to have opposite signs, which is a sufficient condition to ensure that $\psi_{j,1} < \psi_{j,2} < \psi_{j,3}$ or $\psi_{j,3} < \psi_{j,2} < \psi_{j,1}$.

Based on these considerations, we assign $\boldsymbol{\alpha}_j$ and $\boldsymbol{\delta}_j$ a joint prior that is a mixture of two truncated multivariate Gaussian distributions with density

$$p(\boldsymbol{\alpha}_j, \boldsymbol{\delta}_j) = \frac{1}{32\pi^2\omega^2\kappa^2} \exp\left\{-\frac{1}{2}\left(\frac{1}{\omega^2}\boldsymbol{\alpha}_j'\boldsymbol{\alpha}_j + \frac{1}{\kappa^2}(\boldsymbol{\delta}_j - \boldsymbol{\mu})'(\boldsymbol{\delta}_j - \boldsymbol{\mu})\right)\right\}\mathbf{1}(\alpha_{j,1} > 0, \alpha_{j,2} < 0)$$
$$+ \frac{1}{32\pi^2\omega^2\kappa^2} \exp\left\{-\frac{1}{2}\left(\frac{1}{\omega^2}\boldsymbol{\alpha}_j'\boldsymbol{\alpha}_j + \frac{1}{\kappa^2}(\boldsymbol{\delta}_j + \boldsymbol{\mu})'(\boldsymbol{\delta}_j + \boldsymbol{\mu})\right)\right\}\mathbf{1}(\alpha_{j,1} < 0, \alpha_{j,2} > 0). \quad (4)$$

Under this prior, the joint density of $\boldsymbol{\alpha}$ satisfies the required sign constraints, but the marginal distributions of both $\alpha_{j,1}$ and $\alpha_{j,2}$ correspond to zero-mean univariate Gaussian distributions with variance ω^2 and, therefore, have full support over the real line. Furthermore, note that this prior satisfies $p(\boldsymbol{\alpha}_j, \boldsymbol{\delta}_j) = p(-\boldsymbol{\alpha}_j, -\boldsymbol{\delta}_j)$, making the prior invariant to reflections of the latent space. This property enables us to address identifiability issues related to reflections of the latent space as a post-processing step (please see Section 2.2).

To select the hyperparameters $\boldsymbol{\mu}$, ω and κ, we focus our attention on the implied prior on the parameter $\theta_{i,j} = P(y_{i,j} = 1 \mid \beta_i, \alpha_{j,1}, \delta_{j,1}, \alpha_{j,2}, \delta_{j,2})$, which is interpretable and comparable across various model formulations. Here, we again follow the literature and target a prior distribution for $\theta_{i,j}$ that is bimodal, placing most of its probability around 0 and 1 (see, e.g., Paganin *et al.* 2023; Spirling and Quinn 2010; Yu and Rodriguez 2021). In practical terms, this assumption means that most of the time, voters are fairly certain about whether they support or oppose a particular measure. Furthermore, because most measures tend to pass, we expect the distribution of θ_i to place slightly more probability on values close to 1. While there are many priors on $\boldsymbol{\mu}$, ω, and κ that would satisfy these requirements, we set $\boldsymbol{\mu} = (-2, 10)'$, $\omega^2 = 25$, and $\kappa^2 = 10$ in the illustrations we discuss in this paper and study the sensitivity of the results to moderate changes in the hyperparameters. Please see Section 4 of the Supplementary Material. Figure 2 presents a histogram for $\theta_{i,j}$ based on 10,000 draws from our chosen prior distribution, and compares it again a second histogram for the same parameter under the model used in Duck-Mayr and Montgomery (2023). It is worthwhile noting that the prior for $\theta_{i,j}$ under both model is very similar, enabling some of the comparisons we present in Section 3.

(a) Under our probit unfolding model   (b) Under the GGUM of Duck-Mayr and Montgomery (2022)
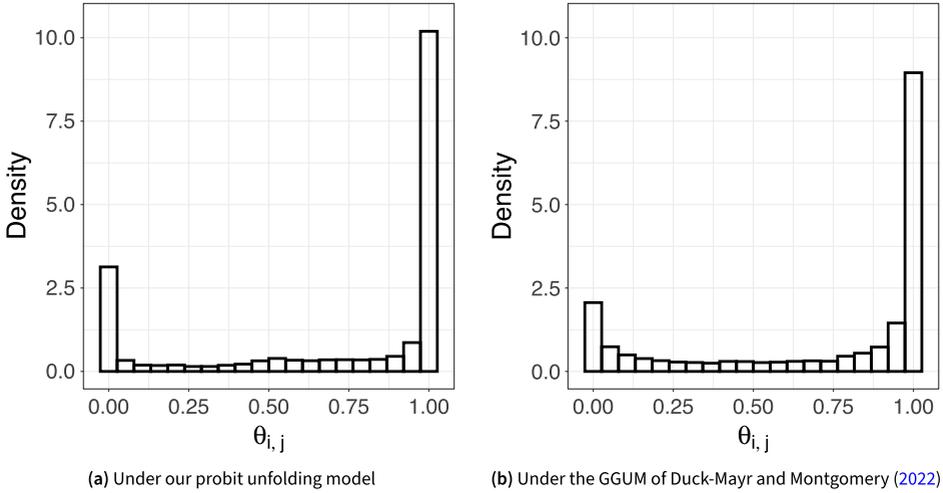
Figure 2. Histograms for 10,000 draws of the implied prior distribution on $\theta_{i,j}$ for our probit unfolding model under a prior with $\boldsymbol{\mu} = (-2, 10)'$, $\omega^2 = 25$, and $\kappa^2 = 10$ compared against the implied prior for the same parameter under the model used in Duck-Mayr and Montgomery (2023).

2.2. Identifiability

Like other spatial voting models, the parameters associated with the latent space are identified in the likelihood only up to an affine transformation. However, the choice of a standard Gaussian prior for the ideal points $\beta_1, \ldots, \beta_I$ means that the posterior is invariant to shifts and rescalings of the latent space and, therefore, weakly identifiable. On the other hand, we address invariance to reflections by fixing the sign of the ideal point of one particular legislator. For example, in the case of the U.S. House of Representatives, we fix the sign of the ideal point of the whip of the Republican party to be positive.

2.3. Computation

We explore the posterior distribution using a Markov chain Monte Carlo algorithm that relies on a data augmentation approach similar to that introduced in Albert and Chib (1993). In particular, for each observations $y_{i,j}$, we introduce three auxiliary variables, $y_{i,j,1}^*$, $y_{i,j,2}^*$, and $y_{i,j,3}^*$, which follow a joint multivariate normal distribution on the form

$$\begin{pmatrix} y_{i,j,1}^* \\ y_{i,j,2}^* \\ y_{i,j,3}^* \end{pmatrix} \Bigg| \alpha_{j,1}, \alpha_{j,2}, \delta_{j,1}, \delta_{j,2}, \beta_i \sim N\left(\begin{pmatrix} -\alpha_{j,1}(\beta_i - \delta_{j,1}) \\ 0 \\ -\alpha_{j,2}(\beta_i - \delta_{j,2}) \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right). \tag{5}$$

Conditioned on these auxiliary variables, the unknown model parameters can be sampled from (mixtures of truncated) Gaussian distributions. Similarly, conditioned on the parameters of interest, the full conditional distributions for the auxiliary variables are truncated Gaussians, with the truncation region being determined by the value of $y_{i,j}$. Hence, it is possible to directly sample from all full conditional posterior distributions and there is no need to tune proposal distributions to specific datasets. Further details of the algorithm are presented in Section 1 of the Supplementary Material.

3. Revealed Preferences in the U.S. House of Representative, 1987–2022

We illustrate the performance of our probit unfolding model by analyzing roll-call voting data from the 100th to the 117th U.S. House of Representatives.[1] We exclude from the analysis legislators who were

---

[1]The replication code and data are available at https://doi.org/10.7910/DVN/SVBF5T (Lei and Rodriguez 2024).

absent for more than 40% of the vote, as well as all unanimous votes. Then, we treat any remaining missing votes as missing completely at random.

We compare our model against IDEAL (Jackman 2001), as well as the BGGUM of Duck-Mayr and Montgomery (2023). Posterior summaries for our model are based on 20,000 iterations of our algorithm obtained after burning the first 200,000 samples and thinning the next 200,000 by a tenth. For IDEAL, we use the algorithm implemented in the R package MCMCpack, and inferences are based on 20,000 iterations obtained after burning the first 10,000 samples. Finally, computation for the BGGUM relies on the algorithm implemented in the R package bggum, and inferences are based on 20,000 samples obtained after burning the first 5,000 iterations and using the next 5,000 to tune the proposal distributions. Please see Duck-Mayr and Montgomery (2023) for details on the bggum package. Convergence of the various algorithms was checked by monitoring the (unnormalized) joint posterior distribution as well as the ideal points of a few legislators in each House using the multi-chain method of Gelman and Rubin (1992).

We start by comparing the fit of the various models using a blocked version of WAIC (Watanabe and Opper 2010; Watanabe 2013; Gelman, Hwang, and Vehtari 2014). For a given House, the WAIC for model m is given by

$$WAIC(m) = -2 \left[\sum_{i=1}^{I} \log \left(\mathrm{E}_{\text{post}} \left\{ \prod_{j=1}^{J} \theta_{i,j}(m)^{y_{i,j}} \left[1 - \theta_{i,j}(m) \right]^{1-y_{i,j}} \right\} \right) \right.$$
$$\left. - \sum_{i=1}^{I} \mathrm{var}_{\text{post}} \left\{ \sum_{j=1}^{J} \left[y_{i,j} \log \theta_{i,j}(m) + (1-y_{i,j}) \log(1-\theta_{i,j}(m)) \right] \right\} \right], \quad (6)$$

where $\theta_{i,j}(m)$ represents the probability that legislator i votes "Aye" in issue j under model m. Lower values of the WAIC provide evidence in favor of that particular model. Like the Akaike Information Criterion and Bayesian Information Criterion, the WAIC attempts to balance model fit with model complexity. However, unlike these two criteria, the WAIC is appropriate for hierarchical setting and is invariant to reparametrizations of the model.

Figure 3a presents the difference in WAIC scores between the probit unfolding model and IDEAL, and between the probit unfolding model and BGGUM. Note that the curve for IDEAL is negative, which is strong evidence that the probit unfolding model outperforms it. Furthermore, we can see that our probit unfolding model outperforms BGGUM in all but three Houses: the 110th, the 111th, and the 112th. Even in these three cases, the advantage of BGGUM is very small. On the other hand, Figure 3b presents the posterior mean (solid lines) and corresponding 95% credible intervals (shaded region) for the Spearman correlation between the legislators' rankings generated by the probit unfolding model and those generated by IDEAL and BGGUM. The Spearman correlation (which, as the name suggests, lies in the $[-1,1]$ interval) is widely used as a distance between rankings, with values close to 1 indicating that the ranking of the legislators generated by both models are identical (Kumar and Vassilvitskii 2010). As would be expected, the rankings based on our probit unfolding model are more similar to those generated by BGGUM than those generated by IDEAL. However, the rankings of the two unfolding models do seem to differ in important ways between the 104th and the 107th Houses. On the other hand, starting with the 110th House, we see that rankings of the two unfolding models are very close but tend to differ substantially from the rankings generated by IDEAL.

To gain additional insight into the behavior of the models, we investigate in more detail the estimates of the rankings for the 103rd, 107th, and 116th Houses. The graphs on the left column of Figure 4 compare the posterior median ranks generated by the probit unfolding model against those generated by IDEAL, while the right column compares the probit unfolding model's ranks against BGGUM's. In line with the results from Figure 3b, we see that all three models tend to yield very similar rankings during the 103rd House. The main outlier is Representative William Lacy Clay Jr. of Missouri. In this case, both unfolding models categorize Clay as more liberal than IDEAL. On the other hand, for the 116th House, the rankings generated by the two unfolding models are very similar, but quite different

(a) WAIC relative to PUM

(b) Spearman Correlation relative to the probit unfolding model
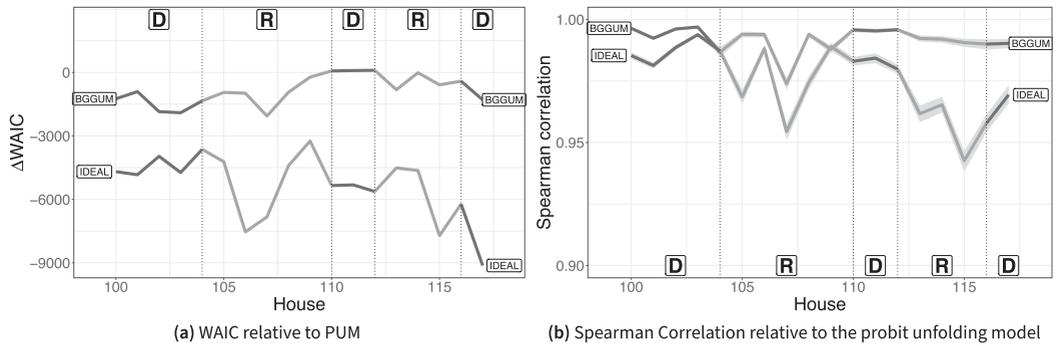
Figure 3. Left panel: Difference in WAIC scores between the probit unfolding model and IDEAL, ($WAIC(\text{PUM}) - WAIC(\text{IDEAL})$), and between the probit unfolding model and BGGUM, ($WAIC(\text{PUM}) - WAIC(\text{BGGUM})$). Right panel: Spearman correlation between the legislators' rankings generated by the probit unfolding model and those generated by either IDEAL or BGGUM.

from those generated by IDEAL. In particular, there are large differences on the rankings of many Democratic legislators, especially those of the so-called "Squad" (Representatives Alexandria Ocasio-Cortez of New York, Ilhan Omar of Minnesota, Ayanna Pressley of Massachusetts, and Rashida Tlaib of Michigan). While both unfolding models rank the members of the Squad as being among the most liberal in the Democratic caucus, IDEAL ranks them as centrists. Based on what we know about the Squad, the rankings generated by the unfolding models seem much more appropriate. The members of the Squad were all elected for the first time in 2018 with support by the Justice Democrats political action committee, and are widely known for advocating progressive policies such as Medicare for All, the Green New Deal, and full student loan debt cancellation. IDEAL tends to rank them as centrists during the 116th House because they often clashed with the Democratic leadership and selectively voted against their own party, often in conjunction with Republican legislators largely understood to be on the extreme of their own party (Lewis 2019a, 2019b). There are also some important differences in the rankings of some Republican legislators during the 116th House, particularly Justin Amash of Michigan and Matt Gaetz of Florida, which are ranked as extremes by the unfolding models. Similar to the Squad on the Democratic side, these legislators are well-known members of the more extreme wing of the Republican party and often voted with extreme Democrats and against their party on a number of key issues.

The previous results suggest that revealed preferences estimated by unfolding models are generally better than those from IDEAL at capturing the ideological leanings of legislators, especially for more extreme ones. However, they do not provide much insight into the differences between the two unfolding models. To address this question, we focus now on the 107th House, which is the one for which Figure 3b shows the largest difference in rankings between the probit unfolding model and BGGUM. In the case of Democratic legislators, both unfolding models mostly agree with each other and produce rankings for the most extreme Democrats that are quite different from IDEAL. Nonetheless, the probit unfolding model and BGGUM yield relative large differences in the rankings of Rick Larsen from Washington, David Wu from Oregon, Jerry Costello from Illinois, Dennis Moore from Kansas, Gary Condit from California, Collin Peterson from Minnesota, and Gene Taylor from Mississippi. In all seven cases, the probit unfolding model ranks these legislators as more centrist than BGGUM. Reviewing their records, it is worthwhile noting that Larsen, Wu, and Moore were all members of the centrist New Democrat Coalition, while Condit, Peterson, and Taylor were all members of the also centrist Blue Dog Coalition (co-founded by Peterson). This suggests that the rankings generated by the probit unfolding model might be more appropriate than those of BGGUM in this case. On the other hand, Figure 4d highlights, among others, the 16 Republican legislators with the largest differences in median ranks. Of these, Jim Ramstad from Minnesota, Frank Lobiondo from New Jersey, Phil English from Pennsylvania, Paul Gillmor from Ohio, Jerry Weller from Illinois, and Melissa Hart from Pennsylvania were all members of the moderate Republican Main Street Partnership and consistently ranked highly
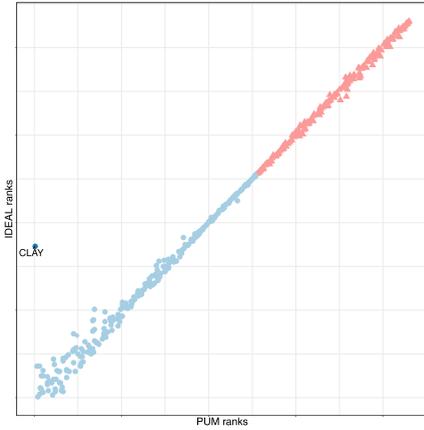
(a) 103rd House: Probit unfolding model vs. IDEAL
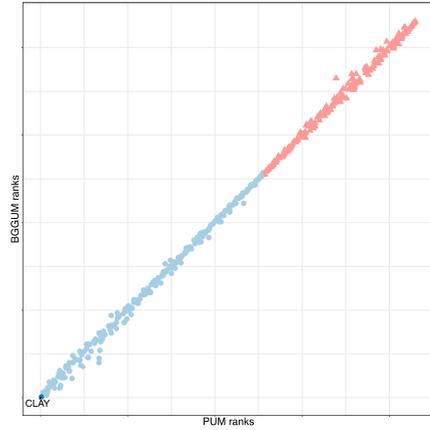
(b) 103rd House: Probit unfolding model vs. BGGUM
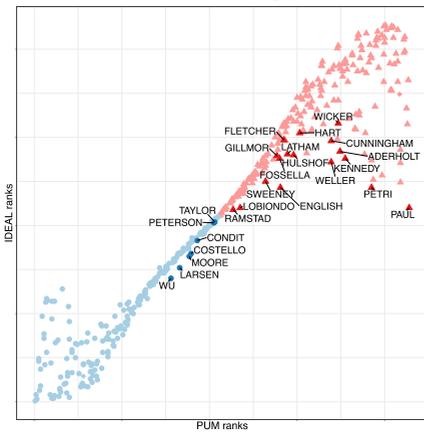
(c) 107th House: Probit unfolding model vs. IDEAL
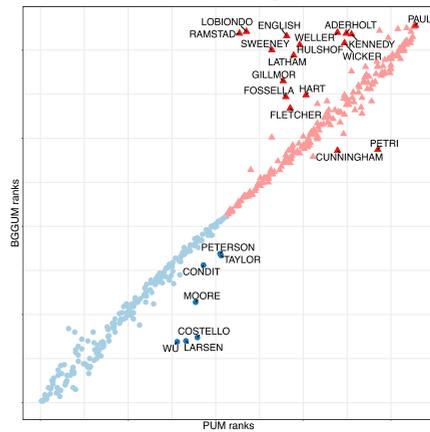
(d) 107th House: Probit unfolding model vs. BGGUM
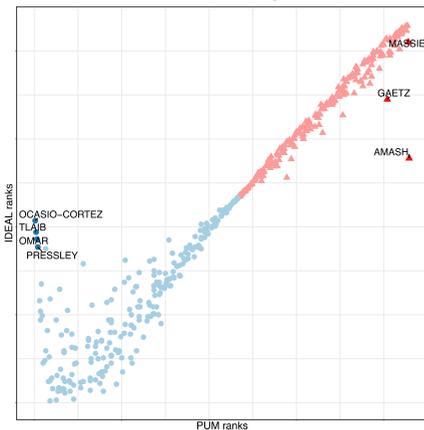
(e) 116th House: Probit unfolding model vs. IDEAL
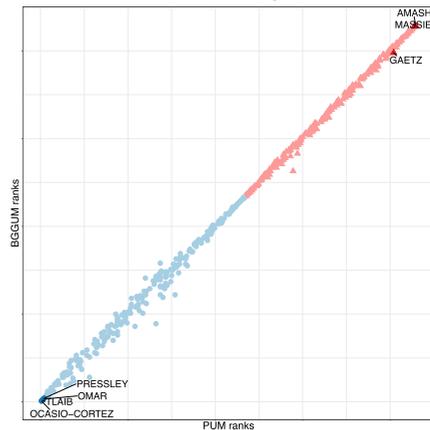
(f) 116th House: Probit unfolding model vs. BGGUM

Figure 4. Comparison of the posterior median ranks of legislators across IDEAL, BGGUM, and our probit unfolding model in selected Houses. Democrats are shown with blue triangles, Republicans are shown with red triangles, and independents are shown with a rhombus and the color of the party that they caucus with.

on bipartisanship indexes. Similarly, John Sweeney from New York, Kenny Hulshof from Missouri, Vito Fossella from New York, and Mark Kennedy from Minnesota had relatively low rankings among Republicans (signaling more moderate ideologies) in Govtrack scores (GovTrack.us 2013). On the other hand, Duke Cunningham (R CA-51), one of the two in the list whose rank under the probit unfolding model is more conservative than under BGGUM, was widely considered an ardent conservative. Again, this suggests that the rankings produced by the probit unfolding model might be a more accurate reflection of legislator's ideology.

We further explore the differences between our profit unfolding model and BGGUM during the 107th House by computing vote-specific WAIC scores. The associated score for vote j under model m is defined analogously to (1):

$$WAIC_j(m) = -2\left[\log\left(E_{post}\left\{\prod_{i=1}^{I} \theta_{i,j}(m)^{y_{i,j}}\left[1 - \theta_{i,j}(m)\right]^{1-y_{i,j}}\right\}\right)\right.$$
$$\left. - var_{post}\left\{\sum_{i=1}^{I}\left[y_{i,j}\log\theta_{i,j}(m) + (1 - y_{i,j})\log(1 - \theta_{i,j}(m))\right]\right\}\right].$$

These scores allow us to identify which votes are better explained by each of the two unfolding model (in terms of complexity-adjusted fit).

Figure 5 shows the difference between vote-specific WAIC scores under BGGUM and the probit unfolding model. About 64% of the votes during the 107th House are better explained by our probit unfolding model. We also see a few votes where the difference of WAIC scores is large in either direction. In particular, there are two votes in which BGGUM substantially outperforms the probit unfolding model. Both of these correspond to Approval of the Journal (ApJ) votes. Article I, section 5 of the U.S. Constitution requires that the House keeps a journal of its proceedings, which is a summary of the day's actions. The Speaker is responsible for examining and approving the Journal of the previous day. Following the announcement of approval by the Speaker, any House member may demand a vote on the question of the Speaker's approval. Journal-approval record votes rarely fail (of 472 held between 1991 and 2016, none failed), but various party and political considerations play a role in the decision to call for such a vote (Hudiburg 2018). Our reading of the literature suggests that voting patterns on Approval of the Journal votes are very idiosyncratic. In particular, Patty (2010) notes that votes on the Journal's approval are just as frequently requested by the majority party as by members of the minority party and that votes recorded on days on which a vote was also recorded on the House Journal are more likely to be close and more likely to be party-line votes than those recorded on other days. Hence, we do not see the fact that BGGUM seems to explain these votes better as strong evidence in its favor. On the other hand, the five highlighted votes that are better explained by the probit unfolding model are all substantive votes with response functions that are estimated to be non-monotonic by both models but where the ranks associated with the probit unfolding model are more consistent with an ends-against-the-middle voting pattern.

The differences in the rankings of legislators we just discussed seem to be driven by substantial differences in the underlying estimates of the ideal points. To illustrate this, we present in Figure 6 the posterior mean of the ratio of the range of the ideal points of Republican and Democratic legislators for each of the three models under consideration. Note that these ratios are invariant to affine transformations of the policy space, and are therefore identifiable and can be compared over time and across models. Generally speaking, we again see that the estimates for the probit unfolding model and for BGGUM are very similar to each other during the first half of the period under study, and quite different from those obtained through IDEAL. The difference is particularly striking in the 100th, 101st, 102nd, and 103rd Houses, where the unfolding models suggest that the ideological spread of Democrats is much larger than that of Republicans, and in the 105th, 106th, 107th, and 115th, where the estimates suggest the opposite. The behavior during the 115th and 116th Houses is particularly interesting. In the 115th House, our model agrees with BGGUM in estimating similar spreads for both parties. This is in contrast to IDEAL, which estimates the spread of Democrats as almost twice as big
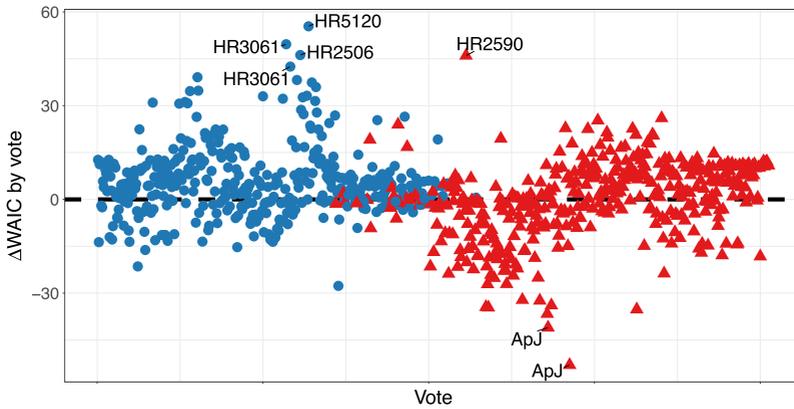
Figure 5. Difference in vote-specific WAIC scores between the probit unfolding model and BGGUM, (WAIC(BGGUM) − WAIC(PUM)), for the 107th House. Note that the way the difference is being computed here is the opposite to the way in which it was computed in Figure 3a. Votes are ordered according to the absolute difference between Republican votes and Democratic votes. Blue circles indicate votes in which the voting Democrats' proportion of "Ayes" is greater than the voting Republicans' proportion of "Ayes," whereas red triangles indicate votes in which the reverse happened.
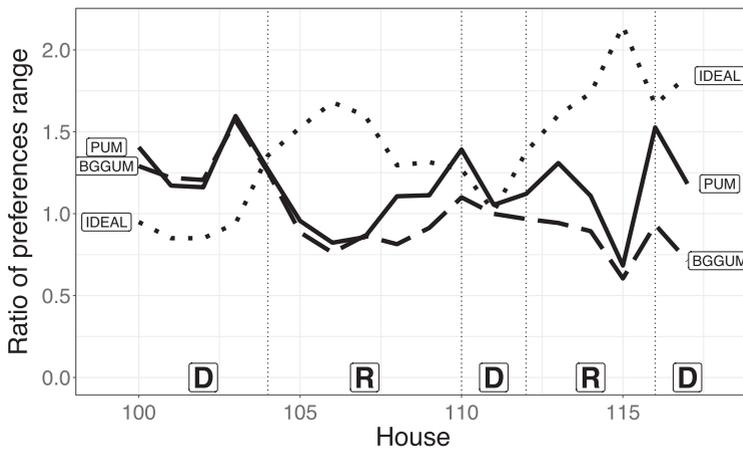

Figure 6. Posterior mean of the ratio of Democrats' range over Republicans' range across the various Houses. The solid line corresponds to the probit unfolding model, the dotted line to IDEAL, and the dashed line to BGGUM.

as that of Republicans. However, in the 116th House (when Democrats regained control of the House), our model closely agrees with IDEAL in estimating a broader spread for Democrats. Section 3 of the Supplementary Material presents alternative versions of Figure 6 based on the standard deviation and the interquartile range of the estimated ideal points. The general patterns are very similar no matter what measure of dispersion is used, but some of the details do differ. For example, the differences between IDEAL and the unfolding models between the 110th and 114th House are less pronounced under the interquartile range and the standard deviation.

Moving on now to a detailed analysis of the 116th House, Figure 7 shows posterior summaries for the distribution of $\alpha_1, \ldots, \alpha_J$. The posterior distribution for 49.9% of the α_j's in this House clearly lie on the lower right quadrant (where $\alpha_{j,1} > 0$ and $\alpha_{j,2} < 0$), while the remaining 51.1% clearly lies on the upper left quadrant. The structure of this posterior distribution provides support for our choice of a prior that allows for α_j to lie on either of the two quadrants, as opposed to a prior that concentrates on just one. On the other hand, Figure 8 shows estimates of the response functions for three different votes. These estimates demonstrate that the probit unfolding model is able to recover both monotonic and

(a) Posterior means

(b) Posterior probability $\alpha_{j,1} > 0$ and $\alpha_{j,2} < 0$
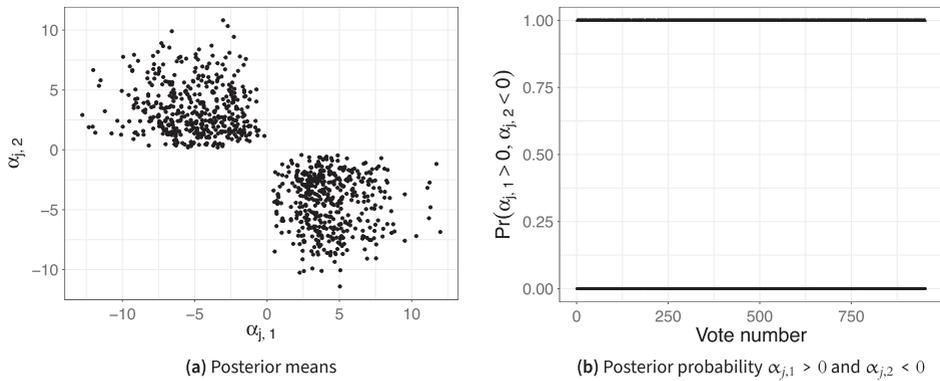
Figure 7. Posterior summaries of 's α_j in the probit unfolding model for the 116th House.


(a) HRES5

(b) HR21

(c) HRES6
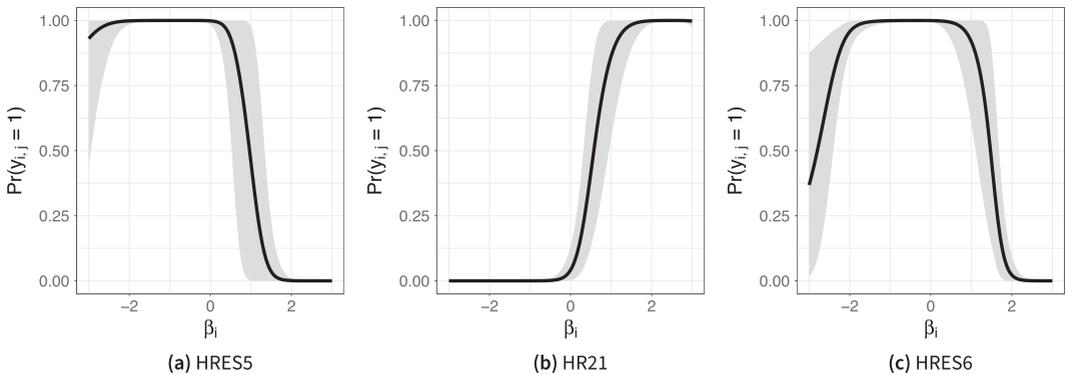
Figure 8. Plots displaying various response curves based on the posterior means of $\alpha_{j,1}, \alpha_{j,2}, \delta_{j,1}$, and $\delta_{j,2}$ from the probit unfolding model for the 116th House. Here, the vote number refers to the clerk's roll-call vote number. Shaded areas correspond to 95% pointwise posterior credible intervals.

non-monotonic response functions from observed data. The left panel of Figure 8 shows the response function for the first in a series of votes on HRES5, which provided for consideration of HRES6 (rules of the House of Representatives for the 116th Congress) and HR21 (appropriations for the FY 2019). This response function is monotonically decreasing, which is to be expected given that the 116th House had a Democratic majority. Indeed, recall from Section 2.2 that the direction of latent space was identified by making the sign of the ideal point of the Republican whip positive. This means that bills with a monotonically decreasing response function are bills favored by Democrats and not by Republicans. Accordingly, all 230 Democrats present voted in favor of HRES5, while all 197 Republicans present voted against it. Similarly, the center panel of Figure 8 shows the response function for the first vote on HR21, a procedural vote proposed by Republicans. The response function in this case is increasing, which agrees with the fact that all Democrats present voted against it while all Republicans voted in favor. Finally, the right panel of Figure 8 shows an example of a non-monotonic response function, corresponding to the first vote taken on HRES6 (rules of the House of Representatives for the 116th Congress). In this case the voting record is mixed: most Democrats and three Republicans voted in favor, while most Republicans and three Democrats voted against the rules. The disagreements within each party are, however, qualitatively different. The three Republican dissenters (Representatives Brian Fitzpatrick, John Katko, and Tom Reed) are considered by most observers as moderates and have often expressed interest in bipartisanship. On the other hand, two of the Democratic dissenters (Representatives Alexandria Ocasio-Cortez and Ro Khanna) are widely considered among the most liberal legislators in the House. The shape of the response function reflects this: the probability of a

positive vote on this bill is high for most Democrats and for the more centrists Republicans, and lower for the more extreme legislators of either party.

4. Dynamic Unfolding Models

The analysis in the previous section relies on fitting the model described in Section 2 independently for each House. While straightforward, this approach does not allow us to fully assess how individual members' preferences evolve over time. In this section, we consider an extension of the probit unfolding model that allows for a longitudinal analyses of member's preferences and extends the approach introduced in Martin and Quinn (2002) to situations in which allowing for non-monotonic response functions might be desirable.

In a manner similar to Section 2, let $y_{i,j,t}$ be the vote of member $i = 1, \ldots, I$ on issue $j = 1, \ldots, J_t$ considered on period $t = 1, \ldots, T$. As before, we postulate that voters make decisions based on three random utility functions,

$$U_{N^-}(\beta_{i,t}, \psi_{j,t,1}) = -(\beta_{i,t} - \psi_{j,t,1})^2 + \epsilon_{i,j,t,1},$$
$$U_Y(\beta_{i,t}, \psi_{j,t,2}) = -(\beta_{i,t} - \psi_{j,t,2})^2 + \epsilon_{i,j,t,2},$$
$$U_{N^+}(\beta_{i,t}, \psi_{j,t,1}) = -(\beta_{i,t} - \psi_{j,t,3})^2 + \epsilon_{i,j,t,3},$$

where $\epsilon_{i,j,t,1}$, $\epsilon_{i,j,t,2}$, and $\epsilon_{i,j,3}$ are independent and identically distributed Gaussian shocks, so that

$$P(y_{i,j,t} = 1 \mid \beta_{i,t}, \alpha_{j,t,1}, \delta_{j,t,1}, \alpha_{j,t,2}, \delta_{j,t,2}) =$$
$$\int_{-\infty}^{\alpha_{j,t,1}(\beta_{i,t} - \delta_{j,t,1})} \int_{-\infty}^{\alpha_{j,t,2}(\beta_{i,t} - \delta_{j,t,2})} \frac{1}{2\sqrt{3}\pi} \exp\left\{-\frac{1}{3}(z_1^2 - z_1 z_2 + z_2^2)\right\} \mathrm{d}z_1 \mathrm{d}z_2.$$

We also continue treating the issue-specific parameters $(\boldsymbol{\alpha}_{j,t}, \boldsymbol{\delta}_{j,t})$ as independent and identically distributed for all j and t, and assign them the prior in Equation (4). The key difference with Section 2 lies on how we model the ideal points $\beta_{i,t}$. Rather than using completely independent priors for all i and t, we treat the vectors $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_I$ as independent, but assign each vector $\boldsymbol{\beta}_i = (\beta_{i,1}, \ldots, \beta_{i,T})'$ a joint Gaussian prior with mean $\mathbf{0}$ and covariance matrix

$$\boldsymbol{\Omega}(\rho) = \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{pmatrix},$$

where the parameter $0 \le \rho \le 1$ is unknown and needs to be learned from the data.

There are several ways in which this joint prior can be motivated. For example, this prior can be obtained as the finite-dimensional marginal of a zero-mean Gaussian process with an exponential covariance function. Alternatively, it can be motivated as a realization of a first-order stationary autoregressive process with autocorrelation ρ and with a standard normal distribution as its stationary distribution. Both of these motivations make it clear that for any t, the marginal distribution of $\beta_{i,t}$ is the same distribution used in Section 2. Furthermore, the parameter ρ controls how much information is borrowed over time; $\rho = 1$ corresponds to a model in which preferences are assumed to be constant over time, whereas $\rho = 0$ simply leads back to fitting independent probit unfolding models for each t. Consistent with Martin and Quinn (2002), in the context of the application described in the next section, we assume that preferences evolve relatively slowly and assign ρ a normal prior with mean $\eta = 0.9$ and standard deviation $\lambda = 0.04$, which is truncated to the interval $[0, 1]$. We also conduct a sensitivity analysis to ascertain the impact of this choice on the inferences we draw from the model. Please see Section 5 of the Supplementary Material for further details.

4.1. Computation

The computational approach described in Section 2.3 can be adapted to the dynamic unfolding model. Most of the steps in the Markov chain Monte Carlo algorithm remain the same, with the main differences being in the sampling of the vector $\boldsymbol{\beta}_i$ and the need for an additional step related to sampling the correlation ρ.

Conditioned on latent vectors $\boldsymbol{y}_{i,j,t}^* = (y_{i,j,t,1}^*, y_{i,j,t,2}^*, y_{i,j,t,3}^*)'$ defined similarly to Equation (5), the posterior distribution for the vector $\boldsymbol{\beta}_i$ reduces to a multivariate normal distribution with a sparse precision matrix. Because of this, sampling from this joint distribution can be done quite efficiently, even for large T, by either exploiting algorithms for sparse linear algebra, or by building a forward-backward algorithm similar to the one employed in Martin and Quinn (2002).[2] As for sampling the autocorrelation ρ, we rely on a Metropolis–Hasting algorithm that uses a Gaussian random walk on $\Upsilon = \log(\rho/(1 - \rho))$ as the proposal distribution. The variance of the random walk is tuned to target a 40% acceptance rate. This is the only step in the algorithm that relies on this type of approach, making implementation relatively straightforward. Details of this algorithm can be found in Section 2 of the Supplementary Material, and code implementing the algorithm is available at https://github.com/rayleigh/probit_unfolding_model/.

5. Revealed Preferences in the U.S. Supreme Court, 1937–2021

We use the dynamic model from Section 4 to examine the voting record of U.S. Supreme Court (SCOTUS) justices between 1937 and 2021. We compare our results to those generated by the dynamic factor model described in Martin and Quinn (2002), which extends IDEAL to dynamic settings and has become the de-facto gold standard method for measuring justice's preferences.

The data we analyze are available at https://mqscores.lsa.umich.edu/. As in Martin and Quinn (2002), a justice's vote is encoded as 1 if they voted to reverse a lower court's decision, and 0 if they voted to affirm that decision. Our inferences are based on 40,000 samples obtained after burning the first 500,000 iterations of our Markov chain Monte Carlo algorithm and thinning the next 400,000 every 10th observation. Results for the model from Martin and Quinn (2002) (MQ in the sequel) were obtained using the R package MCMCpack. We use 20,000 samples from the associated Markov chain Monte Carlo algorithm, obtained after burning the first 40,000 draws, from four chains. This gives us a total of 80,000 samples.

Figure 9a presents the difference in WAIC scores between MQ and the dynamic unfolding model for each of the terms over consideration (recall Equation (1), and note that the difference is being computed in the opposite way to Figure 3a). Under this metric, the dynamic unfolding model seems to consistently outperform MQ. The exceptions are 1965, 1967, and 1975–1979, where MQ seems to dominate. Complementing these results, Figure 9b presents the posterior mean and corresponding 95% credible intervals for the Spearman correlation between the justices' rankings generated by the dynamic probit unfolding model and MQ. We can see that, in spite of the probit unfolding model dominating MQ in terms of WAIC scores, their rankings of the justices agree considerably, especially over the last 40 years. Nevertheless, there are two periods during which the models drastically disagree: 1949–1952 and 1967–1970.

To be concise, we focus the rest of our discussion on the 1949–1952 period, which is also a period when the WAIC score greatly favors our dynamic unfolding model. Between 1949 and 1952, SCOTUS was composed of five appointees of president Franklin D. Roosevelt (Justices Felix Frankfurter, Robert Jackson, Hugo Black, Willian O. Douglas, and Stanley F. Reed), joined by four more recent appointees of Harry S. Truman (Justices Harold H. Burton, Fred M. Vinson, Tom C. Clark, and Sherman Minton, with Clark and Minton serving their first term in 1949). Figure 10 shows the posterior mean of the ideal points for these Justices under the dynamic unfolding model (left column) and MQ (right

---

[2]The alternative, sampling from the full conditional of each $\beta_{i,t}$, leads to an algorithm that mixes extremely poorly and takes a very long time to explore the posterior distribution. This is one of the challenges with trying to extend the approach of Duck-Mayr and Montgomery (2023) to dynamic settings.
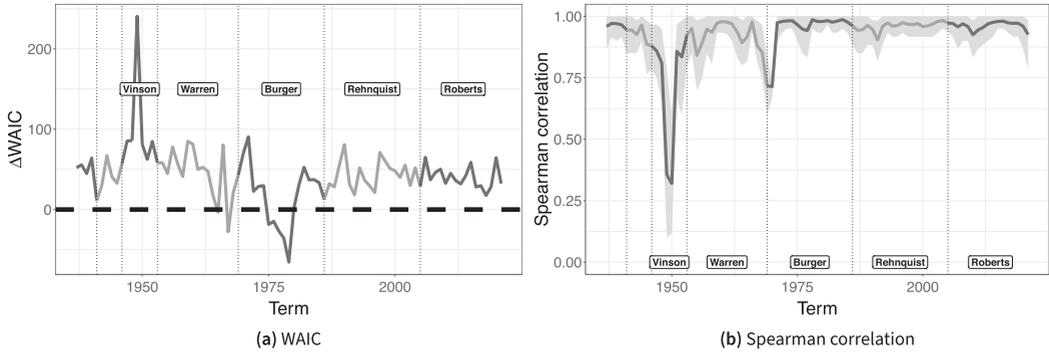
Figure 9. Left panel: Difference in WAIC scores between MQ and the dynamic unfolding model ($WAIC(\text{MQ}) - WAIC(\text{DPUM})$). Note that the way the difference is being computed here is the opposite to the way in which it was computed in Figure 3a. Right panel: Posterior mean (solid line) and corresponding 95% credible intervals (shaded region) for the Spearman correlation between the justices' rankings generated by the dynamic unfolding model and MQ.
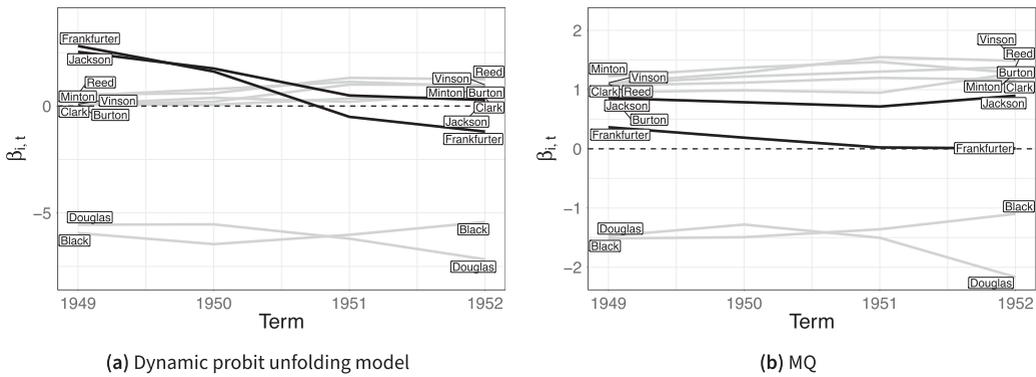

Figure 10. Posterior means of the ideal points for SCOTUS Justices active during 1949–1952 terms under the dynamic unfolding model (left column) and MQ (right column).

column). We can see that the different rankings are driven by substantial differences in the estimates of Justices Frankfurter and Jackson's preferences: MQ places these two justices as centrists, while our dynamic unfolding model places them as the two most conservative members of the court in the 1949–1950 terms. We argue that the characterization coming out of the dynamic unfolding model is in closer agreement with the accepted understanding of these Justices' ideological leanings. Indeed, Justice Frankfurter represents a bit of an enigma for historians and legal scholars, but most have come to see him as the leader of the conservative faction of the Supreme Court (see, e.g., Eisler and Eisler 1993), to which Jackson (a frequent ally of Frankfurter) also belonged. During his long career, first as a professor at Harvard and then as a Justice, Frankfurter was a staunch supporter of judicial restraint. During the 1920s and early 1930s, liberals embraced judicial restraint as a way to check the power of the conservative justices that dominated the Supreme Court, placing Frankfurter firmly on the left wing of the judicial community. However, as Roosevelt set to remake the Supreme Court, most liberals (but not Frankfurter) abandoned judicial restraint and embraced judicial activism. This, along with his difficult relationship with other members of the court, explains how one of the minds behind Roosevelt's New Deal legislation and the creation of the American Civil Liberties Union came to be considered a stalwart of the conservative wing of the Supreme Court.

6. Discussion

This paper introduced a new class of unfolding models for binary preference data that can be motivated from first principles as a spatial voting model. We also consider extensions to dynamic models that allow

the preferences of legislators to evolve over time. A key feature of this class of models is that it allows for non-monotonic response functions, which can arise in practice when legislators at the extremes of the political spectrum vote together against the center. Our extensive evaluations on voting data from the U.S. House of Representatives and the U.S. Supreme Court indicate that the model substantially outperforms both traditional scaling models and alternative unfolding models that had been previously introduced in the literature.

Our model is slightly more flexible than BGGUM, as it does not constrain the discrimination parameters $\alpha_{j,1}$ and $\alpha_{j,2}$ (recall Equation (1) and the associated discussion). Of course, having a model with more parameters is not always advantageous. However, the values of the WAIC in Sections 3 and 5 indicate that, in the vast majority of cases, the additional flexibility of the probit unfolding model improves the overall fit enough to compensate for the slight increase in model complexity. Furthermore, our results suggest that, while both BGGUM and our model sometimes misclassify relative centrists as extremists (or vice versa!), BGGUM tends to do so much more often. In fact, the work presented in this manuscript bears on the question of how to estimate *ideological* rankings from voting data. While the term *ideology* has a long and varied history of usage in scholarship (see, e.g., Gerring 1997), it is most often used to refer to specific policy views and preferences held by individuals, either "an underlying philosophy on which all specific political views are based" (Jessee 2012, 17) or a belief system that includes a wide range of opinions consistently held (Converse 1964). However, the term is commonly operationalized in terms of revealed preferences and ideal points, as estimated by spatial voting models (e.g., Poole and Rosenthal 2006). How faithful this operationalization is to the more etimological definitions is open to debate. One key observation from our studies is that our unfolding model not only leads to superior complexity-adjusted fit measures such as the WAIC, but it also yields vote-based estimates of preferences that seem to better match what, on the basis of the public record, would seem to be the true philosophy/belief system of legislators and justices.

Moving forward, there are several extensions of the framework introduced in this paper that we would like to to pursue. One refers to the class of link functions used to define the model. Instead of using a probit link, we might instead be interested in a logit link, which is more robust to outlier votes. Computation in this case is potentially more challenging, but the algorithm described in this paper could be adapted by using a mixture approximation for the Gumbel distribution (see, e.g., Frühwirth-Schnatter and Frühwirth 2012). A second extension refers to the use of higher-dimensional policy spaces. Right now, all unfolding models that we are aware of assume that the underlying latent space is unidimensional. Our framework provides a natural setting in which to build higher-dimensional models in a principled fashion. Finally, we are interested in extending our framework to general multinomial ordinal observations. Such extension can be achieved by introducing pairs of issue-specific coordinates for each level of the observed categorical variables.

References

Albert, J. H., and S. Chib. 1993. "Bayesian Analysis of Binary and Polychotomous Response Data." *Journal of the American Statistical Association* 88 (422): 669–679.

Clark, J. H. 2012. "Examining Parties as Procedural Cartels: Evidence from the US States." *Legislative Studies Quarterly* 37 (4): 491–507.

Clinton, J. D. 2012. "Using Roll Call Estimates to Test Models of Politics." *Annual Review of Political Science* 15: 79–99.

Converse, P. E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, edited by D. E. Apter, 206–261. London: Free Press of Glencoe.

Davis, O. A., M. J. Hinich, and P. C. Ordeshook. 1970. "An Expository Development of a Mathematical Model of the Electoral Process." *American Political Science Review* 64 (2): 426–448.

De La Torre, J., S. Stark, and O. S. Chernyshenko. 2006. "Markov Chain Monte Carlo Estimation of Item Parameters for the Generalized Graded Unfolding Model." *Applied Psychological Measurement* 30 (3): 216–232.

Duck-Mayr, J., and J. Montgomery. 2023. "Ends against the Middle: Measuring Latent Traits When Opposites Respond the Same Way for Antithetical Reasons." *Political Analysis* 31 (4): 606–625.

Eisler, K. I., and K. T. Eisler. 1993. *A Justice for All: William J. Brennan, Jr., and the Decisions That Transformed America*. New York: Simon & Schuster.

Enelow, J. M., and M. J. Hinich. 1984. *The Spatial Theory of Voting: An Introduction*. New York: CUP Archive.

Frühwirth-Schnatter, S., and R. Frühwirth. 2012. "Bayesian Inference in the Multinomial Logit Model." *Austrian Journal of Statistics* 41 (1): 27–43. http://doi.org/10.17713/ajs.v41i1.186

Gelman, A., J. Hwang, and A. Vehtari. 2014. "Understanding Predictive Information Criteria for Bayesian Models." *Statistics and Computing* 24 (6): 997–1016.

Gelman, A., and D. B. Rubin. 1992. "Inference from Iterative Simulation Using Multiple Sequences." *Statistical Science* 7 (4 [November]): 457–472. https://doi.org/10.1214/ss/1177011136

Gerring, J. 1997. "Ideology: A Definitional Analysis." *Political Research Quarterly* 50 (4): 957–994.

GovTrack.us. 2013. "Ideology Analysis of Members of Congress." https://www.govtrack.us/about/analysis.

Hudiburg, J. A. 2018. *The House Journal: Origin, Purpose, and Approval.* https://crsreports.congress.gov/product/pdf/R/R45209.

Hug, S. 2010. "Selection Effects in Roll Call Votes." *British Journal of Political Science* 40 (1): 225–235.

Jackman, S. 2001. "Multidimensional Analysis of Roll Call Data via Bayesian Simulation: Identification, Estimation, Inference, and Model Checking." *Political Analysis* 9 (3 [January]): 227–241. https://doi.org/10.1093/polana/9.3.227

Jenkins, S. 2006. "The Impact of Party and Ideology on Roll-Call Voting in State Legislatures." *Legislative Studies Quarterly* 31 (2): 235–257.

Jessee, S. A. 2012. *Ideology and Spatial Voting in American Elections*. Cambridge: Cambridge University Press.

Jessee, S. A., and S. M. Theriault. 2014. "The Two Faces of Congressional Roll-Call Voting." *Party Politics* 20 (6): 836–848.

Kumar, R., and S. Vassilvitskii. 2010. "Generalized Distances between Rankings." In *Proceedings of the 19th International Conference on World Wide Web*, 571–580. New York: Association for Computing Machinery.

Lei, R., and A. Rodriguez. 2024. *Replication Data for: A Novel Class of Unfolding Models for Binary Preference Data.* https://doi.org/10.7910/DVN/SVBF5T

Lewis, J. 2019a. *Why Are Ocasio-Cortez, Omar, Pressley, and Talib Estimated to Be Moderates by NOMINATE?* https://voteview.com/articles/Ocasio-Cortez_Omar_Pressley_Tlaib.

Lewis, J. 2019b. *Why Is Alexandria Ocasio-Cortez Estimated to Be a Moderate by NOMINATE?* https://voteview.com/articles/ocasio_cortez.

Lofland, C. L., A. Rodriguez, and S. Moser. 2017. "Assessing Differences in Legislators' Revealed Preferences: A Case Study on the 107th US Senate." *Annals of Applied Statistics* 11 (1): 456–479.

Luque, C., and J. Sosa. 2023. "A Bayesian Spatial Voting Model to Characterize the Legislative Behavior of the Colombian Senate 2010–2014." *Journal of Applied Statistics* 50 (16): 3362–3383.

Martin, A. D., and K. M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999." *Political Analysis* 10 (2): 134–153. https://doi.org/10.1093/pan/10.2.134

McFadden, D. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, edited by P. Zarembka, 105–142. New York: Academic Press.

Moser, S., A. Rodríguez, and C. L. Lofland. 2021. "Multiple Ideal Points: Revealed Preferences in Different Domains." *Political Analysis* 29 (2 [April]): 139–166. https://doi.org/10.1017/pan.2020.21

Paganin, S., C. Paciorek, C. Wehrhahn, A. Rodriguez, S. Rabe-Hesketh, and P. de Valpine. 2023. "Computational Methods for Bayesian Semiparametric Item Response Theory Models." *Journal of Educational and Behavioral Statistics* 48: 147–188.

Patty, J. W. 2010. "Dilatory or Anticipatory? Voting on the Journal in the House of Representatives." *Public Choice* 143 (1–2): 121–133.

Poole, K., and H. Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science* 29 (2): 357–384.

Poole, K., and H. Rosenthal. 2006. *Ideology and Congress: A Political Economic History of Roll Call Voting.* 2nd ed. New Brunswick: Routledge.

Roberts, J. M. 2007. "The Statistical Analysis of Roll-Call Data: A Cautionary Tale." *Legislative Studies Quarterly* 32 (3): 341–360.

Roberts, J. S., J. R. Donoghue, and J. E. Laughlin. 2000. "A General Item Response Theory Model for Unfolding Unidimensional Polytomous Responses." *Applied Psychological Measurement* 24 (1): 3–32.

Rodriguez, A., and S. Moser. 2015. "Measuring and Accounting for Strategic Abstentions in the US Senate, 1989–2012." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 64 (5): 779–797.

Schickler, E. 2000. "Institutional Change in the House of Representatives, 1867–1998: A Test of Partisan and Ideological Power Balance Models." *American Political Science Review* 94 (2): 269–288.

Schwindt-Bayer, L. A., and R. Corbetta. 2004. "Gender Turnover and Roll-Call Voting in the US House of Representatives." *Legislative Studies Quarterly* 29 (2): 215–229.

Shor, B., and N. McCarty. 2011. "The Ideological Mapping of American Legislatures." *American Political Science Review* 105 (3): 530–551.

Spirling, A., and K. Quinn. 2010. "Identifying Intraparty Voting Blocs in the UK House of Commons." *Journal of the American Statistical Association* 105 (490): 447–457.

Tendeiro, J. N., and S. Castro-Alvarez. 2019. "GGUM: An R Package for Fitting the Generalized Graded Unfolding Model." *Applied Psychological Measurement* 43 (2): 172–173.

Tu, N., B. Zhang, L. Angrave, and T. Sun. 2021. "Bmggum: An R Package for Bayesian Estimation of the Multidimensional Generalized Graded Unfolding Model with Covariates." *Applied Psychological Measurement* 45 (7–8): 553–555.

Wang, W., J. de la Torre, and F. Drasgow. 2015. "MCMC GGUM: A New Computer Program for Estimating Unfolding IRT Models." *Applied Psychological Measurement* 39 (2): 160–161.

Watanabe, S. 2013. "A Widely Applicable Bayesian Information Criterion." *Journal of Machine Learning Research* 14 (1): 867–897.

Watanabe, S., and M. Opper. 2010. "Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory." *Journal of Machine Learning Research* 11 (12): 3571–3591.

Yu, X., and A. Rodriguez. 2021. "Spatial Voting Models in Circular Spaces: A Case Study of the US House of Representatives." *Annals of Applied Statistics* 15 (4): 1897–1922.