

# SPATIAL VOTING MODELS IN CIRCULAR SPACES: A CASE STUDY OF THE U.S. HOUSE OF REPRESENTATIVES

BY XINGCHEN YU<sup>1</sup> AND ABEL RODRÍGUEZ<sup>2</sup>

<sup>1</sup>*Department of Statistics, University of California, Santa Cruz, [xyu26@ucsc.edu](mailto:xyu26@ucsc.edu)*

<sup>2</sup>*Department of Statistics, University of Washington, [abelrod@uw.edu](mailto:abelrod@uw.edu)*

The use of spatial models for inferring members’ preferences from voting data has become widespread in the study of deliberative bodies, such as legislatures. Most established spatial voting models assume that ideal points belong to a Euclidean policy space. However, the geometry of Euclidean spaces (even multidimensional ones) cannot fully accommodate situations in which members at the opposite ends of the ideological spectrum reveal similar preferences by voting together against the rest of the legislature. This kind of voting behavior can arise, for example, when extreme conservatives oppose a measure because they see it as being too costly, while extreme liberals oppose it for not going far enough for them. This paper introduces a new class of spatial voting models in which preferences live in a circular policy space. Such geometry for the latent space is motivated by both theoretical (the so-called “horseshoe theory” of political thinking) and empirical (goodness of fit) considerations. Furthermore, the circular model is flexible and can approximate the one-dimensional version of the Euclidean voting model when the data supports it. We apply our circular model to roll-call voting data from the U.S. Congress between 1988 and 2019 and demonstrate that, starting with the 112th House of Representatives, circular policy spaces consistently provide a better explanation of legislators’s behavior than Euclidean ones and that legislators’s rankings, generated through the use of the circular geometry, tend to be more consistent with those implied by their stated policy positions.

**1. Introduction.** Spatial voting models (Davis, Hinich and Ordeshook (1970), Enelow and Hinich (1984), Poole and Rosenthal (1985), Heckman and Snyder Jr. (1996), Jackman (2001), Clinton, Jackman and Rivers (2004), Clinton and Jackman (2009)) are widely used to estimate the preferences of legislators from roll-call voting records and have become an invaluable tool in the study of legislatures and other deliberative bodies. Spatial voting models aim to scale binary and polychotomous responses into a continuous (potentially multidimensional) linear scale and are related to traditional statistical tools for dimensionality reduction, such as principal components and factor analysis. Spatial voting models are also intimately related to item response theory (IRT) models (Fox (2010), Wilson and De Boeck (2004), van der Linden and Hambleton (1997)) which are widely used in educational testing to estimate abilities, attitudes or other unobserved features of respondents. In the context of voting data, the latent space on which the responses are scaled is referred to as the *policy space*, while the latent traits are referred to as the *ideal points* of the legislators.

Traditional spatial voting models rely on latent spaces endowed with Euclidean geometries, and, therefore, tend to work best in political systems in which the parties are relatively unified. Because of this, they struggle to explain voting patterns of legislatures in two-party systems in which parties are “fractious.” In this kind of setting, it is common to see legislators that most observers would consider to be at opposite ends of the ideological spectrum vote together. Spirling and McLean (2007) and Spirling and Quinn (2010) consider one example,

---

Received September 2020; revised January 2021.

*Key words and phrases.* Spatial voting model, roll-call votes, U.S. Congress, factor model, spherical geometry, geodesic Hamiltonian Monte Carlo.

namely, the first Blair government (1997–2001) in the United Kingdom. This government represented an uneasy alliance between a leadership that “had actively abandoned the tenants of socialist policy making and that had received a landslide mandate to rule” and a “historically and openly recalcitrant tranche of ‘Old’ Labour legislators, dismissive of the modernizing project in its entirety” (Spirling and Quinn (2010)). In the United States the conservative revolt led by the Tea Party movement during the 2010 election (Karpowitz et al. (2011), Arceneaux and Nicholson (2012), Skocpol and Williamson (2016)) and the recent rise of the Justice Democrats during the 2018 election (Lewis (2019a, 2019b)) represent two more examples. Traditional spatial models fail in this setting where the “extremes vote together” because, under the Euclidean geometry, the “rebels” who sometimes vote with the opposition must necessarily be placed somewhere in the middle of the scale. Neither increasing the dimensionality of the latent space nor performing linear transformations of the latent space can address this issue (see Section 2).

In order to gain insights into legislatures in which the extremes vote together, Spirling and Quinn (2010) proposed a Bayesian nonparametric mixture model that identifies voting blocks within the U.K. House of Common. In a similar spirit, Guimerà and Sales-Pardo (2011) and Crane (2017) developed random partition models for studying the voting record of the U.S. Supreme Court. While this kind of clustering models can provide valuable insights into the functioning of a deliberative body, they do not yield the kind of fine-grained ranking that has made spatial voting models so useful in practice. There is also an interesting literature focusing on the effect of the underlying utility functions on the estimates of the ideal points in Euclidean settings. For example, Carroll et al. (2013) describe a model in which the form of the utility function (quadratic or Gaussian on the Euclidean distance between points) is estimated from the data and conclude that extreme legislators are generally more sensitive to policy changes than their more centrally located counterparts, while Humphreys and Laver (2010) discuss the use of the “city-block” (i.e.,  $L^1$ ) instead of  $L^2$  distances in multidimensional spatial models and Eguia (2013) considers the general case of Minkowski (i.e.,  $L^q$ ) distances, where  $q$  is a parameter that is to be estimated from the data. More recently, Duck-Mayr and Montgomery (2020) develop a model with nonmonotone utility functions to explain the phenomenon of extremes voting together against the center. As an alternative, this paper proposes a novel spatial voting model that relies on a circular policy space and develops Bayesian inference procedures for it. An appealing feature of our model is that it can approximate the IDEAL model (Jackman (2001), Clinton, Jackman and Rivers (2004)), when the variance of the ideal points is small, yielding a measure of the “circularity” of the policy space that can be very helpful in characterizing historical voting patterns.

This paper is motivated by the study of contemporary voting patterns in the U.S. House of Representatives, where the “extremes voting together” phenomenon has become increasingly common. For that reason our focus is on one-dimensional models voting models. Indeed, although there is evidence that, historically, voting patterns in the U.S. Congress were well explained by two-dimensional (Euclidean) models, it is widely accepted that, starting in the mid-to-late 1970s, the U.S. Congress has become increasingly unidimensional (e.g., see McCarty, Poole and Rosenthal (2016), Hare and Poole (2014), Poole and Rosenthal (1991, 1997), Moser, Rodríguez and Lofland (2021)). In one-dimensional policy spaces, the ideal points generated by spatial voting models are often interpreted as capturing the ideology of the legislator on a liberal-conservative scale (e.g., see Poole and Rosenthal (1985) and Jessee (2012)) with the ranking of the legislators in this scale typically becoming a key metric of interest.

The idea that circular policy spaces might be appropriate for representing political preferences dates back at least to Weisberg (1974), who provides a number of examples and notes that “circular shapes may be expected for alliance structures and for vote coalitions

where extremists of the left and right coalesce for particular purposes.” Our model can also be understood as operationalizing the so-called “Horseshoe Theory” (Pierre (2002), Taylor ((2006), page 118)), which asserts that the far-left and the far-right are closer to each other than they are to the political center, in an analogous way to how the opposite ends of a horse-shoe are close to each other.

There is a rich literature on methods for embedding data into data-driven, non-Euclidean manifolds. Examples include autoencoders (e.g., Kramer (1991) and Kingma et al. (2019)), locally-linear embeddings (Roweis and Saul (2000)), Isomap (Tenenbaum, De Silva and Langford (2000)), Laplacian eigenmaps (Belkin and Niyogi (2002)), local tangent space alignment (Zhang and Zha (2004)) and Gaussian process latent variable models (Lawrence (2005)). These approaches are very flexible in capturing the geometry of the underlying manifold on which the data lives and can produce very compact representations of high-dimensional data using a very small number of latent dimensions. However, for the most part they have been designed for continuous observations rather than multivariate binary data, making their use in voting data suspect. More importantly, interpreting and quantifying the uncertainty associated with the embeddings generated by most of these methods can be quite difficult. Indeed, note that, when performing nonlinear embeddings, invariance to affine transformations is not enough to ensure identifiability of the latent features. When the goal is prediction, this rarely matters. However, when the goal is scaling categorical data, identifiability becomes a key issue for both interpretation and uncertainty quantification. Our approach is different from the ones listed above in that we consider a fixed but more general manifold on which to project the data so that the model is more flexible than its Euclidean counterpart but identifiability issues can be easily addressed. A similar philosophy has been recently adopted in the network modeling literature; for example, see and Smith, Asta and Calder (2019) and McCormick and Zheng (2015).

**2. A motivating example: Ranking “The Squad”.** The November 2018 midterm election saw the Democratic Party win a new majority in the House of Representatives on the back of a record number of women, young and minority candidates. Particularly notable among them is a group of four new members (Alexandria Ocasio-Cortez of New York, Ilhan Omar of Minnesota, Ayanna Pressley of Massachusetts and Rashida Tlaib of Michigan, all women of color under 50 supported by the Justice Democrats political action committee), who often refer to themselves as “The Squad.” As discussed in Lewis (2019a, 2019b), The Squad is widely understood to belong to the left wing of the Democratic party, supporting policies such as the green new deal, reparations for slavery and abolishing the Immigration and Customs Enforcement Agency. Partly because of their support for these policies, they have shown a willingness to challenge the leadership of their party and to vote against it on some issues.

Table 1 presents the rank order of the members of the Squad on a liberal-conservative scale based on their voting record during the first session of the 116th Congress (extending between January 3, 2019 and January 3, 2020). These rankings, in which lower numbers correspond to more liberal legislators, were obtained by fitting one- and two-dimensional versions of the Euclidean model described in Jackman (2001) and Clinton, Jackman and Rivers (2004) (see also Section 3 below). Counterintuitively, all members of The Squad are ranked toward the center of the political spectrum under both models. Most importantly, note that the addition of a second dimension does not dramatically affect the original surprising conclusion that they all appear to belong to the moderate wing of the Democratic party. As we discussed in the Introduction, this counterintuitive result is a direct consequence of the Euclidean geometry underlying these models: If a legislator votes with the opposite party against the majority of its own, the only possible explanation is that the legislator is a moderate.

TABLE 1

*Median rank of the members of “The Squad” during the first session of the 116th U.S. House of Representatives according to two scaling models: A one-dimensional Euclidean model and a two-dimensional Euclidean model. In the case of the two dimensional model, the ranking provided is along the first (highest variability) dimension of the policy space. Lower numbers for the ranks correspond to more liberal legislators. Numbers in parentheses correspond to 95% credible intervals*

|                         | Rank order     |                |
|-------------------------|----------------|----------------|
|                         | Euclidean (1D) | Euclidean (2D) |
| Pressley (D MA-7)       | 172 (131,196)  | 168 (113,200)  |
| Omar (D MN-5)           | 176 (135,198)  | 160 (98,195)   |
| Tlaib (D MI-13)         | 180 (146,200)  | 169 (108,200)  |
| Ocasio-Cortez (D NY-14) | 203 (185,215)  | 197 (172,213)  |

To further investigate the voting behavior of The Squad, we also fitted the nonparametric mixture model described in [Spirling and Quinn \(2010\)](#) to these data. The model identifies three groups of Democrats that appear to have distinct behavior: a small group of 18 legislators representing some of the districts that were flipped by Democrats during the 2018 election and whose seats are widely understood to be at most risk in the 2020 election (we could call these the *vulnerables*), a medium sized group of 61 legislators that include most of the remaining representatives from flipped districts as well as a number of legislators with relatively short tenures in the House (we could call them the *pragmatists*) and a large group of 155 legislators that includes the leadership as well as most representatives with a long tenure in the House (call them the *establishment*). Interestingly, for our purposes the members of The Squad are not split off into a separate group that includes left-wing activists but are, instead, clustered with the establishment. Note that, because of the structure of the [Spirling and Quinn \(2010\)](#) model, no further rankings of the legislators are possible within each block.

**3. Bayesian spatial voting models with Euclidean policy spaces.** This section briefly reviews the Bayesian model introduced of [Jackman \(2001\)](#) and [Clinton, Jackman and Rivers \(2004\)](#), whose framework serves as motivation for our approach. Let  $y_{i,j} \in \{0, 1\}$  be the vote of legislator  $i = 1, \dots, I$  on question  $j = 1, \dots, J$ , where  $y_{i,j} = 1$  if legislator  $i$  voted “Yea” on question  $j$  and  $y_{i,j} = 0$  if the legislator voted “Nay.” Each legislator is assumed to have a preferred policy outcome, which is represented by a position on a (possibly  $q$ -dimensional) latent policy space,  $\beta_i \in \mathbb{R}^q$ . Similarly, each question that is posed to legislators has associated with it two positions in the same policy space: one associated with a “Yea” answer, denoted  $\psi_j \in \mathbb{R}^q$ , and one associated with a “Nay” answer, denoted  $\zeta_j \in \mathbb{R}^q$ . Then, given these positions, legislators make their choice about how to vote independently from one another and from other questions based on the value of two random utilities (e.g., see [McFadden \(1973\)](#)),

$$(1) \quad U_{\text{Yea}}(\psi_j, \beta_i) = -\|\psi_j - \beta_i\|^2 + \varepsilon_{i,j}, \quad U_{\text{Nay}}(\zeta_j, \beta_i) = -\|\zeta_j - \beta_i\|^2 + v_{i,j},$$

where  $\varepsilon_{i,j}$  and  $v_{i,j}$  represent random shocks to the utilities. Under these assumptions a rational actor will vote “Yea” if and only if  $U_{\text{Yea}}(\psi_j, \beta_i) > U_{\text{Nay}}(\zeta_j, \beta_i)$ . Further assuming that the difference between the random shocks  $v_{i,j}$  and  $\varepsilon_{i,j}$  are independent and follow a distribution with cumulative distribution  $G_j(x) = G(x/\sigma_j)$ , we have

$$(2) \quad \Pr(y_{i,j} = 1 \mid \beta_i, \psi_j, \sigma_j, \zeta_j) = G(\mu_j + \alpha_j^T \beta_i),$$

where  $\alpha_j = 2(\psi_j - \zeta_j)/\sigma_j$  can be interpreted as the discrimination ability of question  $j$  (i.e., its ability to distinguish between “liberals” and “conservatives”) and  $\mu_j = (\zeta_j^T \zeta_j - \psi_j^T \psi_j)/\sigma_j$  controls the baseline probability of an affirmative vote. For computational simplicity it is common to use (hierarchical) Gaussian priors for the  $\mu_j$ s,  $\alpha_j$ s and  $\beta_i$ s so that computation can proceed using well-established Markov chain Monte Carlo algorithms (e.g., see [Bafumi et al. \(2005\)](#)).

Note that (2) corresponds to a generalized linear model with link function  $G$ . In particular, when  $G$  corresponds to the cumulative distribution function of the normal distribution, we obtain a probit model. On the other hand, the case in which  $G$  corresponds to a logistic distribution leads to the a logit model. Furthermore, because both  $\alpha_j$  and  $\beta_i$  are unknown, this is just an instance of a factor analysis model for binary data. Additionally, note that the parameters are not all identifiable in this model. For example, the scale of the link function,  $\sigma_j$ , cannot be separately identified from the scale associated with the policy space. Hence, it is common to set  $\sigma_j = 1$  for all  $j = 1, \dots, J$ . Furthermore, the model is invariant to affine transformations of the latent space. In political science applications this is usually dealt with by either fixing the position of  $q + 1$  legislators (such as the party whips, e.g., see [Clinton, Jackman and Rivers \(2004\)](#)) or by fixing the location and scale of the ideal points, along with constraints on the matrix of discrimination parameters (e.g., see [Geweke and Singleton \(1981\)](#)).

**4. Bayesian spatial voting models with circular policy spaces.** The framework described in Section 3 lends itself naturally to extensions to policy spaces with more general geometric properties. In particular, we can embed the latent positions  $\beta_i$ ,  $\psi_j$  and  $\zeta_j$  on a Riemannian manifold  $\mathcal{D}$  and then replace the Euclidean distance used in the definition of the utility functions in (1) with the geodesic distance  $\rho$  on  $\mathcal{D}$  so that

$$(3) \quad U_{\text{Yea}}(\psi_j, \beta_i) = -\rho(\psi_j, \beta_i)^2 + \varepsilon_{i,j}, \quad U_{\text{Nay}}(\zeta_j, \beta_i) = -\rho(\zeta_j, \beta_i)^2 + v_{i,j}.$$

In this paper we focus on the special case where  $\mathcal{D}$  corresponds to the unit circle so that  $\beta_i, \psi_j, \zeta_j \in [-\pi, \pi]$  can be interpreted as angular positions on the circle, and  $\rho(a, b) = \arccos(\cos(a - b))$  is just the smallest angle separating  $a$  and  $b$ . Because of the conditional independence among votes, this formulation leads to a likelihood function of the form

$$(4) \quad \Pr(Y | \psi, \zeta, \beta) = \prod_{i=1}^I \prod_{j=1}^J [G_{\kappa_j}(e_{i,j}(\psi_j, \zeta_j, \beta_i))]^{y_{i,j}} [1 - G_{\kappa_j}(e_{i,j}(\psi_j, \zeta_j, \beta_i))]^{1-y_{i,j}},$$

where  $Y$  is the  $I \times J$  data matrix with entries  $y_{i,j}$ ,  $\psi = (\psi_1, \dots, \psi_J)^T$ ,  $\zeta = (\zeta_1, \dots, \zeta_J)^T$  and  $\beta = (\beta_1, \dots, \beta_I)^T$  are the vectors of unknown positions for all legislators and questions in the policy space,  $G_{\kappa_j}$  is the cumulative distribution function associated with  $v_{i,j} - \varepsilon_{i,j}$  and

$$e_{i,j}(\psi_j, \zeta_j, \beta_i) = \{\arccos(\cos(\zeta_j - \beta_i))\}^2 - \{\arccos(\cos(\psi_j - \beta_i))\}^2.$$

Figure 1 provides some intuition for the additional flexibility intrinsic to the circular voting model and, in particular, for its ability to accommodate situations in which “the extremes vote together.” The left panel depicts a situation in which the outcome of a vote follows along party lines. This type of situation, in which the “Yea” and “Nay” positions fall close to the center of mass of opposite parties, represents the most typical type of question in most legislatures and is well modeled using traditional Euclidean policy spaces. In particular, it captures situations in which moderates from one party vote with the other party. In contrast, the right panel depicts a situation in which moving the “Yea” and “Nay” positions to the upper and lower poles leads, with the same ideal points as before, to an outcome in which the “extreme” members of each party join forces in voting against the question.



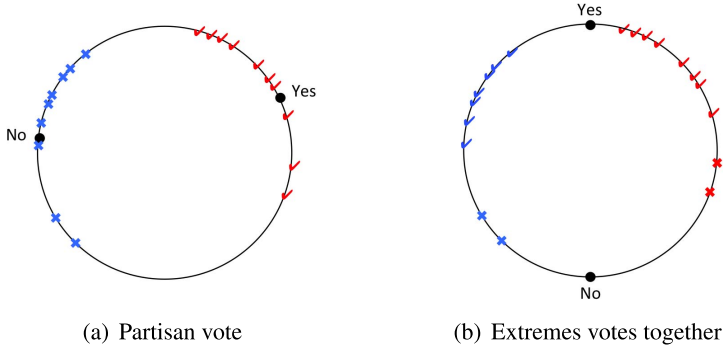


FIG. 1. Two configurations in a circular policy space. Check marks and crosses correspond to the ideal points of legislators voting in favor and against a question and are the same on both panels. Circles correspond to the “Yea” and “Nay” positions for the questions. The left panel, in which the bill positions are located in the upper hemisphere, corresponds to a vote along party lines. The right panel, in which the “Yea” and “Nay” positions fall in the upper and lower poles, corresponds to a question in which the extremes vote together.

**4.1. Link function.** Selecting a link function for the model is nontrivial. Note that, unlike the Euclidean distance in  $\mathbb{R}^q$ , the geodesic distance on the circle takes values in the interval  $[0, \pi]$ . This means that the difference between two squared distances takes values in  $[-\pi^2, \pi^2]$ , and our link function must account for this. We propose to define  $G_{\kappa_j}$  as the cumulative distribution function of a scaled and shifted symmetric beta distribution with density,

$$(5) \quad g_{\kappa_j}(z) = \frac{1}{2\pi^2} \frac{\Gamma(2\kappa_j)}{\Gamma(\kappa_j)\Gamma(\kappa_j)} \left( \frac{\pi^2 + z}{2\pi^2} \right)^{\kappa_j-1} \left( \frac{\pi^2 - z}{2\pi^2} \right)^{\kappa_j-1}, \quad z \in [-\pi^2, \pi^2].$$

The use of this transformed symmetric beta distribution has two advantages in this setting. First, the parameter  $\kappa_j$  has a direct interpretation as a precision parameter. Indeed, the variance of a random variable with density (5) is equal to  $\pi^4/(2\kappa_j + 1)$ . This provides a direct analogy with the scaling parameter  $\sigma_j$  introduced in Section 3. Second, note that

$$(6) \quad \lim_{\kappa_j \rightarrow \infty} \frac{g_{\kappa_j}(z)}{\sqrt{\frac{2\kappa_j+1}{2\pi^5}} \exp\{-\frac{2\kappa_j+1}{\pi^4} z^2\}} = 1,$$

that is, as the concentration parameter increases, the density  $g_{\kappa_j}$  resembles that of a normal distribution with zero mean and variance  $\pi^4/(2\kappa_j + 1)$ . This limit behavior will play an important role when discussing the relationship between our circular model and traditional Euclidean models (see Section 4.4).

**4.2. Identifiability.** As mentioned in the [Introduction](#), the goal of our circular model is scaling rather than prediction. Thus, the identifiability of the latent traits  $\beta_1, \dots, \beta_I$  is crucial. We discuss here the constraints required to make all model parameters identifiable.

To start, notice that the likelihood in equation (4) remains constant if the same shift is applied to all  $\beta_i$ s,  $\psi_j$ s and  $\zeta_j$ s. We address this location invariance through a careful selection of the prior distribution on the  $\beta_i$ s (see Section 4.3). Furthermore, the likelihood also remains constant if any angle is independently increased or decreased by  $2\pi$ . This invariance to “wrappings around the circle” is easily addressed by mapping all angles to the  $[-\pi, \pi]$  interval. Finally, note that the model is invariant to reflections of the policy space, just like the one-dimensional Euclidean model. We address this by fixing the sign of the ideal point of one particular legislator (e.g., the whip of one of the parties).

A key difference between the Euclidean and circular models, however, is that the positions in the circular model are not invariant to changes in scale. As a consequence, the parameters  $\kappa_1, \dots, \kappa_J$  controlling the variance of the link function in (4) are identifiable and can be estimated separately from the  $\beta_i$ s,  $\psi_j$ s and  $\zeta_j$ s. In fact, because the geodesic distance  $\rho$  is bounded, learning  $\kappa_j$ s from the data and allowing them to vary across questions is key to accommodate the full variety of voting behaviors and, in particular, unanimous votes (see Section 6 for some illustrations).

**4.3. Prior distributions.** Since  $\beta_i$ ,  $\zeta_j$  and  $\psi_j$  represent angles, it is natural to use independent von Mises distributions for these parameters:

$$(7) \quad \begin{aligned} \beta_i &| \omega_\beta, \tau_\beta \sim \text{vonMis}(\tau_\beta, \omega_\beta), \\ \psi_j &| \omega_\psi, \tau_\psi \sim \text{vonMis}(\tau_\psi, \omega_\psi), \\ \zeta_j &| \omega_\zeta, \tau_\zeta \sim \text{vonMis}(\tau_\zeta, \omega_\zeta). \end{aligned}$$

A random variable  $Z$  follows a von Mises distributions with mean  $\tau$  and concentration  $\omega$ ,  $Z \sim \text{vonMis}(\tau, \omega)$ , if it has density

$$p(z) = \frac{1}{2\pi I_0(\omega)} \exp\{\omega \cos(z - \tau)\}, \quad z \in [-\pi, \pi],$$

where  $I_k(\omega)$  is the modified Bessel function of order  $k$ . When  $\omega = 0$ , the von Mises distribution becomes the uniform distribution on the circle. On the other hand, as  $\omega$  grows, the distribution behaves as a normal distribution with variance  $1/\omega$ , that is,

$$(8) \quad \lim_{\omega \rightarrow \infty} \frac{\frac{1}{2\pi I_0(\omega)} \exp\{\omega \cos(z - \tau)\}}{\sqrt{\frac{\omega}{2\pi}} \exp\{-\frac{\omega}{2}(z - \mu)^2\}} = 1.$$

In fact, we can think about the von Mises as being the circular analogue of the Gaussian distribution.

To elicit the hyperparameters of the model, we rely on the intuition provided by Figure 1. Since we want the question's positions to potentially be located anywhere on the circle, we set  $\omega_\psi = \omega_\zeta = 0$  (leading, as we mentioned before, to uniform priors on the circle for these two parameters). On the other hand, the ideal points are assigned a zero mean, that is,  $\tau_\beta = 0$  and a nonzero precision, that is,  $\omega_\beta > 0$ . This structure ensures (weak) identifiability of all latent positions to location shifts (recall our discussion from Section 4.2). In particular, we let  $\omega_\beta$  be an exponential hyperprior with mean  $\theta = 10$  so that  $\Pr(-\pi/2 < \beta_i < \pi/2) \approx 0.95$  and perform a sensitivity analyses. This choice for  $\omega_\beta$  reflects our prior belief that a Euclidean model is reasonable in most cases, and, therefore, most ideal points will be concentrated on the upper hemisphere (see discussion in Section 4.4). Finally, we assume that the  $\kappa_j$ s are independent and identically distributed a priori from an exponential prior with mean  $\lambda$  which is, in turn, assigned a (conditionally conjugate) inverse Gamma prior with one degree of freedom and rate parameter 25 (i.e.,  $1/\lambda$  follows an exponential distribution with mean  $\xi = 1/25$ ). Again, we investigate the impact of this choice in our applications through a sensitivity analysis.

**4.4. Relationship with traditional Euclidean models.** Under certain circumstances our spherical model can behave similarly to the probit version of the one-dimensional Euclidean model described in Section 3. To understand this relationship, consider projecting the latent angles that describe the circular model onto the tangent space at 0 (see Figure 2). Two points need to be made about such projection. First, note that, if the latent positions concentrate around the northern pole, the projection of the angles  $\beta_i$ ,  $\psi_j$  and  $\zeta_j$  onto the tangent space

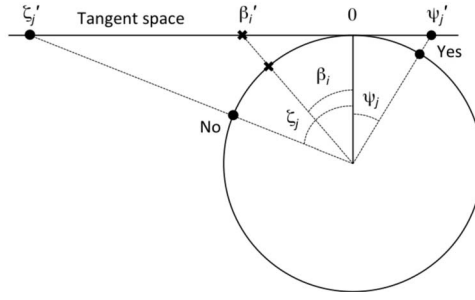


FIG. 2. The circular manifold and its projection on the tangent space at the origin (located in our graphs at the upper pole). The values  $\beta_i$ ,  $\psi_j$  and  $\zeta_j$  correspond to the coordinates in the circular policy space (measured as angles with respect to vertical axis), while the values of  $\beta'_i$ ,  $\psi'_j$  and  $\zeta'_j$  are their projections on the tangent space.

(labeled  $\beta'_i$ ,  $\psi'_j$  and  $\zeta'_j$  in the figure) satisfy  $\beta'_i = \tan \beta_i \approx \beta_i$ ,  $\psi'_j = \tan \psi_j \approx \psi_j$  and  $\zeta'_j = \tan \zeta_j \approx \zeta_j$  for large values of  $\omega_\beta$ . Furthermore, under those circumstances,  $\rho(\psi_j, \beta_i) \approx |\psi_j - \beta_i| \approx |\psi'_j - \beta'_i|$  and  $\rho(\zeta_j, \beta_i) \approx |\zeta_j - \beta_i| \approx |\zeta'_j - \beta'_i|$ , that is, the geodesic distance between the points in the manifold is very close to the Euclidean distance between their projections on the tangent space. Second, recall from (6) that, as  $\kappa_j \rightarrow \infty$ , the link function  $G_{\kappa_j}$  we have chosen for the circular model will resemble the cumulative distribution of the normal distribution with variance  $\frac{\pi^4}{2\kappa_j+1}$ . As a consequence of these two features, if we let both  $\omega_\beta \rightarrow \infty$  and  $\kappa_j \rightarrow \infty$ , while keeping  $\frac{\pi^4 \omega_\beta}{2\kappa_j+1} = 1$ , the value of the likelihood function for the circular model can be expected to be very similar to the value of the likelihood of a one-dimensional Euclidean model with a probit link constructed on the tangent space at 0. Furthermore, under these circumstances the von Mises prior on the circular coordinates maps onto the widely used Gaussian prior on the tangent space (recall equation (8)).

The previous discussion suggests that we can use the variance of the ideal points to measure the level of circularity in the policy space of a given dataset. In particular, small values for this variance indicate that the policy space might be approximately Euclidean. We will make use of this observation in Section 6.3.

Another useful interpretation of our model that arises from this connection is as an interpolator between the one-dimensional and the two-dimensional Euclidean models. Indeed, in addition to the natural geometric argument that arises from embedding the circle into a two-dimensional Euclidean space, we note that the likelihood associated with the 1D Euclidean model involves  $o(I + 2J)$  parameters, the one for the circular model involves  $o(I + 3J)$  parameters and the one for the two-dimensional Euclidean involves  $o(2I + 4J)$ . This means that the circular model provides slightly more degrees of freedom to fit the data than a one-dimensional Euclidean model but less than a two-dimensional Euclidean.

**4.5. Ranking legislators under the circular model.** The unit circle is not endowed with a total order which represents a challenge if our goal is to rank legislators using the latent scale. We get around this issue by unwinding the circle into a traditional linear scale in  $(-\pi, \pi]$ . Breaking the circle at the bottom pole is natural if we consider the fact that the prior on the ideal points is centered at 0 (which corresponds to the middle of this interval) as well as the behavior of the prior when  $\omega_\beta \rightarrow \infty$ .

Unwinding the circle might seem somewhat ad hoc after our heavy emphasis on the circular nature of the policy space. A formal justification is as follows: if the ideal points  $\beta_1, \dots, \beta_I$  all lie on the interval  $(-\pi/2, \pi/2)$  (i.e., the upper semicircle in Figure 1), the ranking of the projection of the ideal points on the tangent space at 0 (given by  $\beta'_i = \tan \beta_i$ , recall Figure 2) is identical to the ranking generated by unwinding the circle (since the tangent is a monotonic



function when restricted to this domain). This scenario (all  $\beta_i$ s in the  $(-\pi/2, \pi/2)$  interval) is not a contrived one: it is an assumption that underlies the construction of our priors and one that seems to be supported in most of the examples discussed in Section 6, including some of those in which there is evidence that the circular model dominates the Euclidean one. Nonetheless, it is important to remember that, if the ideal point of any legislator is very close to either  $-\pi$  or  $\pi$ , the ranks generated by this procedure might be very sensitive to small perturbations and, potentially, misleading.

**5. Computation.** The posterior distribution for the model, which takes the form

$$(9) \quad p(\boldsymbol{\beta}, \boldsymbol{\zeta}, \boldsymbol{\psi}, \boldsymbol{\kappa}, \omega_\beta, \lambda \mid \mathbf{Y}) \\ \propto \left[ \prod_{i=1}^I \prod_{j=1}^J \{G_{\kappa_j}(e_{i,j}(\psi_j, \zeta_j, \beta_i))\}^{y_{i,j}} \{1 - G_{\kappa_j}(e_{i,j}(\psi_j, \zeta_j, \beta_i))\}^{1-y_{i,j}} \right] \\ \times \left[ \prod_{i=1}^I \frac{\exp\{\omega_\beta \cos(\beta_i)\}}{2\pi I_0(\omega_\beta)} \right] \left[ \frac{1}{\theta} \exp\left\{-\frac{\omega_\beta}{\theta}\right\} \right] \left[ \prod_{j=1}^J \frac{1}{\lambda} \exp\left\{-\frac{\kappa_j}{\lambda}\right\} \right] \left[ \frac{1}{\xi \lambda^2} \exp\left\{-\frac{1}{\xi \lambda}\right\} \right],$$

is analytically intractable. Hence, inference for the model parameters is carried out using Markov chain Monte Carlo (MCMC) algorithms.

The algorithm we propose is a hybrid that combines Gibbs sampling, random walk Metropolis–Hastings and Hamiltonian Monte Carlo (HMC) steps to sample from the conditional distributions of each parameter. The simplest steps correspond to sampling the parameters  $\lambda$ ,  $\omega_\beta$  and  $\kappa_1, \dots, \kappa_J$ . More specifically, we sample  $\lambda$  from its inverse-Gamma full conditional posterior distribution and sample  $\omega_\beta$  as well as each of the  $\kappa_j$ s using random walk Metropolis–Hastings with log-Gaussian proposals. The variance of these proposals are tuned so that the acceptance rate is roughly 40%. On the other hand, for sampling the latent positions we employ the geodesic Hamiltonian Monte Carlo (GHMC) algorithm described in [Byrne and Girolami \(2013\)](#). This class of algorithms provides a scalable and efficient way to obtain samples from target distributions defined on manifolds that can be embedded in a Euclidean space. In our particular setting, GHMC samplers work by mapping angles into a two-dimensional coordinate system and operating over the associated Hausdorff measure.

As an example, consider the step associated with updating each of the  $\beta_i$ s. From equation (9), the density (with respect to the Lebesgue measure on  $[-\pi, \pi]$ ) of its full conditional distribution takes the form

$$p(\beta_i \mid \dots) \propto \exp\{\omega_\beta \cos(\beta_i)\} \\ \times \prod_{j=1}^J \{G_{\kappa_j}(\{\arccos(\cos(\zeta_j - \beta_i))\}^2 - \{\arccos(\cos(\psi_j - \beta_i))\}^2)\}^{y_{i,j}} \\ \times \{1 - G_{\kappa_j}(\{\arccos(\cos(\zeta_j - \beta_i))\}^2 - \{\arccos(\cos(\psi_j - \beta_i))\}^2)\}^{1-y_{i,j}},$$

while the density of the associated Hausdorff measure in  $\mathbb{R}^2$  is given by

$$p(\mathbf{x}_{\beta_i} \mid \dots) \propto \exp\{\boldsymbol{\eta}_\beta^T \mathbf{x}_{\beta_i}\} \prod_{j=1}^J \{G_{\kappa_j}(\{\arccos(\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})\}^2 - \{\arccos(\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})\}^2)\}^{y_{i,j}} \\ \times \{1 - G_{\kappa_j}(\{\arccos(\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})\}^2 - \{\arccos(\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})\}^2)\}^{1-y_{i,j}}, \quad \mathbf{x}_{\beta_i}^T \mathbf{x}_{\beta_i} = 1,$$

where  $\boldsymbol{\eta}_\beta^T = (\omega_\beta, 0)$ ,  $\mathbf{z}_{\psi_j}^T = (\cos \psi_j, \sin \psi_j)$ ,  $\mathbf{z}_{\zeta_j}^T = (\cos \zeta_j, \sin \zeta_j)$  and the mapping from  $\beta_i$  to  $\mathbf{x}_{\beta_i}$  is  $\mathbf{x}_{\beta_i}^T = (\cos \beta_i, \sin \beta_i)^T$ . Given tuning parameters  $\varepsilon$  (the step size) and  $L$  (the number of steps), the GHMC sampler takes the form:

1. Map the current value of the chain,  $\beta_i^{(c)}$  onto the embedding space  $\mathbb{R}^2$  by setting  $\mathbf{x}_{\beta_i}^{(c)} = (\cos \beta_i^{(c)}, \sin \beta_i^{(c)})^T$ , and initialize  $\mathbf{x}_{\beta_i} = \mathbf{x}_{\beta_i}^{(c)}$ .
2. Initialize the auxiliary momentum variable  $\phi$  by sampling  $\phi \sim N(0, \mathbf{I}_2)$ .
3. Project the momentum onto the tangent space at  $\mathbf{x}_{\beta_i}$  by setting  $\phi \leftarrow (\mathbf{I}_2 - \mathbf{x}_{\beta_i} \mathbf{x}_{\beta_i}^T) \phi$ , and then set  $\phi^{(c)} = \phi$ .
4. For each of the  $L$  leap steps:
  - (a) Update the momentum by setting  $\phi \leftarrow \phi + \frac{\varepsilon}{2} \nabla \log p_{\mathcal{H}}(\mathbf{x}_{\beta_i} | \dots)$ .
  - (b) Project the momentum onto the tangent space at  $\mathbf{x}_{\beta_i}$  by setting  $\phi \leftarrow (\mathbf{I}_2 - \mathbf{x}_{\beta_i} \mathbf{x}_{\beta_i}^T) \phi$ , and then set the angular velocity of the geodesic flow  $v = \|\phi\|$ .
  - (c) Update  $\mathbf{x}_{\beta_i}$  and  $\phi$  jointly according to the geodesic flow with step size of  $\varepsilon$ ,
 
$$\mathbf{x}_{\beta_i} \leftarrow \mathbf{x}_{\beta_i} \cos(v\varepsilon) + \frac{\phi}{v} \sin(v\varepsilon), \quad \phi \leftarrow \phi \cos(v\varepsilon) - v \mathbf{x}_{\beta_i} \sin(v\varepsilon).$$
  - (d) Update  $\phi \leftarrow \phi + \frac{\varepsilon}{2} \nabla \log p_{\mathcal{H}}(\mathbf{x}_{\beta_i} | \dots)$ .
  - (e) Project the momentum onto the tangent space at  $\mathbf{x}_{\beta_i}$  by setting  $\phi \leftarrow (\mathbf{I}_2 - \mathbf{x}_{\beta_i} \mathbf{x}_{\beta_i}^T) \phi$ .
5. Set the proposed values as  $\mathbf{x}_{\beta_i}^{(p)} = \mathbf{x}_{\beta_i}$ ,  $\phi^{(p)} = \phi$  and  $\beta^{(p)} = \arctan2(x_{\beta_i,2}, x_{\beta_i,1})$ . The proposed value  $\beta^{(p)}$  is accepted with probability

$$\min \left\{ 1, \frac{p_{\mathcal{H}}(\mathbf{x}_{\beta_i}^{(p)} | \dots) \exp\{-\frac{1}{2}[\phi^{(p)}]^T \phi^{(p)}\}}{p_{\mathcal{H}}(\mathbf{x}_{\beta_i}^{(c)} | \dots) \exp\{-\frac{1}{2}[\phi^{(c)}]^T \phi^{(c)}\}} \right\}.$$

Detailed expressions for the Hausdorff measures associated with the full conditional distributions of the  $\beta_i$ 's,  $\psi_j$ 's and  $\zeta_j$ 's, as well as their gradients, can be seen in the [Appendix](#). In our implementation of the algorithm, we periodically vary the value of the tuning parameters  $\varepsilon$  and  $L$  by randomly sampling them from predetermined distributions. These changes are done independently of the current value of the parameter, thereby preserving the Markovian structure of the algorithm. This approach, sometimes called “jittering” in the literature (e.g., see [Gelman et al. \(\(2014\), page 306\)](#)), greatly improve the mixing of the algorithm in our experiments. The specific range in which  $\varepsilon$  and  $L$  move for each parameter and dataset is selected to target an average acceptance probability between 60% and 90% ([Beskos et al. \(2013\)](#), [Betancourt, Byrne and Girolami \(2014\)](#)). The source code for the methods implemented in this paper is available in the Supplementary Materials ([Yu and Rodriguez \(2021b\)](#)) and the algorithm has also been implemented in an R package available at <https://github.com/Xingchen-Yu/SLFM1D>.

**6. Circular voting in the modern U.S. Congress.** In this section we analyze roll-call voting data from the modern U.S. House of Representatives (1987–2019). There is broad agreement in the literature that, during this period, the U.S. Congress can be considered to be unidimensional ([McCarty, Poole and Rosenthal \(2016\)](#), [Hare and Poole \(2014\)](#), [Poole and Rosenthal \(1991, 1997\)](#), [Moser, Rodriguez and Lofland \(2021\)](#)). Hence, our focus is on understanding the geometry of this intrinsically unidimensional spaces. We first present legislator-level results for two specific Houses (the 116th and the 112th) and then show a longitudinal analysis of chamber-level summaries covering the 100th to the 116th Houses. In all of these analyses, the number of leaps used in the HMC steps is randomly selected from a discrete uniform distribution between one and 10 every 50 samples. Similarly, the leap sizes are drawn from uniform distribution on (0.01, 0.04) or (0.005, 0.04) for each  $\beta_i$  and from a uniform distribution on (0.01, 0.105) for each  $\zeta_j$  and  $\psi_j$ . All inference presented

in this section are based on 20,000 samples obtained after convergence of the Markov chain Monte Carlo algorithm. The length of the burn in period varied between 20,000 and 80,000 iterations, depending on the dataset, with a median around 30,000. Convergence was checked by monitoring the value of the log-likelihood function, both through visual inspection of the trace plot and by comparing multiple chains using the procedure in Gelman and Rubin (1992). Details on the convergence and mixing analysis can be seen in Section 1 of the Supplementary Materials (Yu and Rodriguez (2021a)). Roll-call data for the U.S. House of Representatives was obtained from <https://voteview.com/>. For each House, we excluded legislators who are absent for more than 40% of the votes. Missing values were treated as if missing completely at random. While this assumption is not completely accurate (e.g., see Rodríguez and Moser (2015)), it is commonly made in most applied settings, and we do not expect it to dramatically affect our analyses. In addition to the analyses of the U.S. Congress, we present in this section, Section 3 of the Supplementary Materials (Yu and Rodriguez (2021a)) includes results of a simulation study that corroborate our observations about the performance of the method.

6.1. *The Squad, revisited.* First, we revisit the voting record of the first session of the 116th Congress discussed in Section 2. Table 2 reports the posterior median rank order and associated 95% credible intervals for the members of The Squad according to our circular model. The difference between these and those we reported in Table 1 is dramatic with the circular model clearly placing Pressley, Omar, Tlaib and Ocasio-Cortez among the most liberal members of the Democratic party. As we discussed before, such placement agrees much better with expert’s understanding of The Squad’s ideological stances, including the fact that Pressley is perhaps the least liberal of the four.

More generally, Figure 3 compares the rank order of legislators estimated using the one-dimensional Euclidean model to the rank order estimated by the circular model. On the Democratic side we can see some substantial differences in the ranks estimated by the Euclidean model vs. those estimated by the circular model. However, it is clear that the largest differences correspond to the four members of The Squad. In contrast, on the Republican side the ranks estimated by both models are generally in close agreement. The three main exceptions are representatives Justin Amash (MI-3), Thomas Massie (KY-4) and Matt Gaetz (FL-1), who are estimated to be more extreme by the circular model (see also Table 3). An inspection of their record suggests that the ranking generated by the circular model is much more sensible. Consider first Justin Amash and Thomas Massie. Justin Amash is a libertarian-leaning conservative first elected in 2010 as a Republican. He has received high scores from from right-leaning interest groups, such as the Club for Growth, Heritage Action for America and Americans for Prosperity, and praise from conservative think tanks and nonprofit organizations. He was also a founding member of the House Freedom Caucus, a group of hard-line

TABLE 2  
*Median rank of the members of “The Squad” during  
the first session of the 116th U.S. House of  
Representatives according to our circular model.  
Lower numbers for the ranks correspond to more  
liberal legislators. Numbers in parenthesis correspond  
to 95% credible intervals*

|                         | Rank order (Circular) |
|-------------------------|-----------------------|
| Pressley (D MA-7)       | 5 (1,21)              |
| Omar (D MN-5)           | 2 (1,8)               |
| Tlaib (D MI-13)         | 2 (1,9)               |
| Ocasio-Cortez (D NY-14) | 3 (1,11)              |

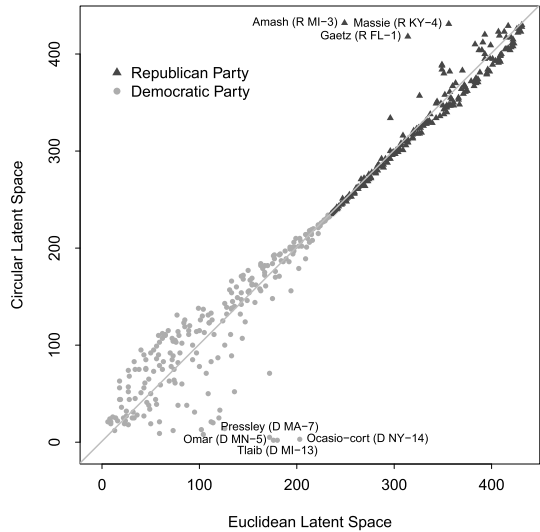


FIG. 3. Posterior median of the rank-order in the first session of the 116th U.S. House of Representatives under the one-dimensional Euclidean (horizontal axis) and the circular (vertical axis) models.

conservative Republicans in the House of Representatives. However, he is also widely known for his contrarian views and for voting with Democrats in certain issues. For example, he was the only Republican to vote against the “In God We Trust” House Resolution passed in November 2011 and the House Resolution supporting the officers and personnel of Immigration and Customs Enforcement (ICE) in July 2018. Furthermore, he cosponsored a bill by Democrat Ayanna Pressley (one of the members of The Squad) that would abolish the death penalty at the federal level. In fact, Amash left the Republican party in July 2019 to become an independent and became the only non-Democrat in the House to vote in favor of an impeachment inquiry into the activities of President Trump and of either of the two articles of impeachment. Thomas Massie is another libertarian-leaning Republican who is often associated with the House Liberty Caucus of Tea Party Republicans. However, he is also know for often being the only member of the House to vote against a number of resolutions. For example, on March 27, 2020, Massie forced the return to Washington of members of the House (who were sheltering in place in the midst of the Covid-19 crisis) by withholding unanimous consent on the passage of the The Coronavirus Aid, Relief and Economic Security (CARES) Act. On the other hand, consider representative Gaetz. Matt Gaetz is a well-known “Trump-ist” firebrand, (in)famous for wearing a gas mask on the House floor during the vote on the

TABLE 3  
*Median rank of three selected Republican legislators during the first session of the 116th U.S. House of Representatives according to two models: a one-dimensional Euclidean voting model and our circular model. Higher numbers for the ranks correspond to more conservative legislators. Numbers in parenthesis correspond to 95% credible intervals*

|                  | Rank order     |               |
|------------------|----------------|---------------|
|                  | Euclidean (1D) | Circular      |
| Amash (R MI-3)   | 249 (244,255)  | 432 (432,432) |
| Massie (R, KY-4) | 356 (334,375)  | 431 (430,431) |
| Gaetz (R FL-1)   | 314 (297,332)  | 418 (412,423) |

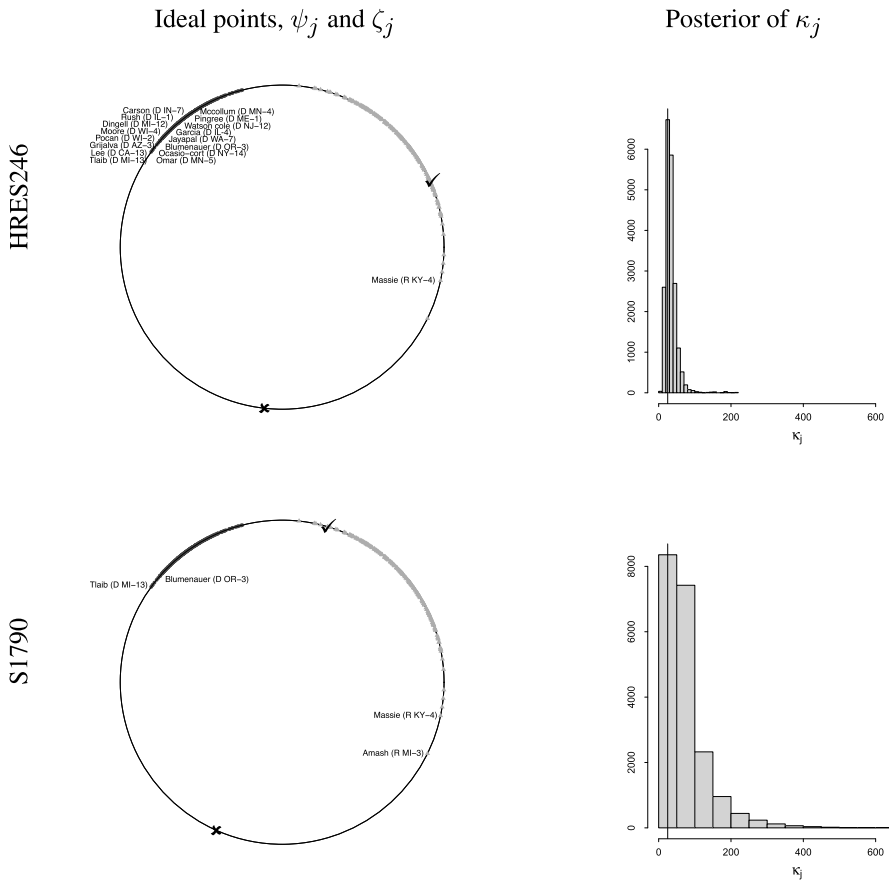


FIG. 4. Two examples of circular voting patterns during the 116th House of Representatives. Graphs on the left column depict the posterior mean ideal point for the legislators (which are the same on both plots), along with the “Yea” and “Nay” positions (represented through a check mark and a cross, respectively). The names in the graphs correspond to the legislators that voted against the measure. The right columns presents a histogram of samples of the posterior distribution of the corresponding  $\kappa_j$ . The vertical line corresponds to  $1/E(1/\lambda \mid \text{data})$ .

first coronavirus response package in March 2020. He ostensibly did so to ridicule the need for such bill, in spite of the fact that it had broad bipartisan support. Even just from this short description, it should be no surprise that our circular model ranks him as being more extreme than the Euclidean model. In fact, in some ways Gaetz makes a better argument for the spherical models than Massie and Amash. (Real) Libertarians are “conservative” in economic issues but “liberal” in social ones, so they make a good intuitive argument for 2D Euclidean models. That is not the case for Gaetz.

To complete this illustration, we provide specific information about various bills in which both circular and Euclidean voting patterns are present. First, Figure 4 provides two examples of circular voting in the 116th House of Representatives. The first one, HRES246, opposed the global boycott, divestment and sanctions movement as well as other efforts targeting Israel. This resolution: (1) urged both sides in the Israel-Palestinian conflict to return to direct negotiations, (2) expressed support for a solution resulting in the State of Israel existing alongside a democratic Palestinian state and (3) reaffirmed the right of U.S. citizens to free speech, including the right to protest or criticize U.S. or foreign government policies. HRES246 was opposed by a group of 16 Democrats (including three members of The Squad) as well as by Representative Massie. Massie opposed the measure because it “calls for Israel to implement a so-called two-state solution. Rather than dictate to Israel what the U.S.

believes is best for Israel, Congress should instead refrain from interfering with Israel's own decisions regarding its foreign and domestic policy." He also stated that he "do[es] not support federal efforts to condemn any type of private boycott, regardless of whether or not a boycott is based upon bad motives. These are matters that Congress should properly leave to the States and to the people to decide." Both of these are traditional "libertarian" arguments. On the other hand, the main driver for Democrats voting against this resolution was support for Palestine. For example, in her floor speech, representative Tlaib invoked her Palestinian grandmother in opposing the resolution, which she said "attempts to delegitimize a certain people's political speech and send a message that our government can and will take action against speech it doesn't like." While there seems to be some ideological common ground between both positions (in particular, a shared desire to limit government impingement on free speech), it is clear that the underlying ideology is completely different, seemingly making this an instantiation of the Horseshoe Theory in the context of the U.S. House of Representatives. The second example in this category is S1790, the National Defense Authorization Act for Fiscal Year 2020. S1790 authorized FY2020 appropriations and set forth policies for Department of Defense (DOD) programs and activities, including military personnel strengths. S1790 was opposed by two Democrats (Tlaib, who is one of the members of The Squad, and Blumenauer) as well as by two Republicans (Massie and Amash, whom we have already discussed). Note that, for both HRES246 and S1790, the "Nay" position is roughly located opposite to the (circular) average of all ideal points, while the "Yea" position is located close to the (circular) average of the ideal points of the legislators that voted in favor of the measure. This matches the intuition captured in Figure 1. Furthermore, the posterior distribution of  $\kappa_j$  indicates moderate to low concentration values for the link function for this kind of votes.

On the other hand, Figure 5 shows two examples of Euclidean voting, HRES5 and HR135. HRES5, which sets forth the rule for consideration of HRES6 (adopting the Rules of the House of Representatives for the 116th Congress), was voted strictly along party lines. We see in this case that the "Yea" and "Nay" positions are located at either side of the parties and that the posterior distribution of  $\kappa_j$  favors relatively large concentration values. This configuration is very similar to the one that is obtained by fitting a Euclidean model to the data. On the other hand, HR135, the Elijah E. Cummings Federal Employee Antidiscrimination Act of 2019, requires each federal agency to establish a model Equal Employment Opportunity Program that is independent of the agency's Human Capital or General Counsel office, and it establishes requirements related to complaints of discrimination and retaliation in the workplace. HR135 was voted unanimously (except for eight abstentions). Note that the "Yea" and "Nay" positions in this case are similar to those estimated for the circular votes, but the corresponding value of  $\kappa_j$  is much higher for the fully unanimous vote.

**6.2. *The conservative revolt of 2010.*** The election in November 2008 of Barack Obama as President of the United States generated a strong conservative backlash that has had a profound impact on U.S. politics in general and on the Republican party in particular (Skocpol and Williamson (2016)). This backlash influenced the results of the 2010 midterm election (Karpowitz et al. (2011)). The 112th Congress had a large Republican majority (in fact, had the largest Republican majority since the 80th Congress in the late 1940s). It was also the first Congress in over 150 years in which the Republican party held the House but not the Senate and the first Congress to begin with the House and the Senate controlled by different parties since the 99th Congress (1985–1987).

Among the 242 Republican legislators elected to the 112th House of Representatives was a large group of insurgent candidates, many of them backed by a loose grassroots coalition ostensibly built on the principles of fiscal responsibility, adherence to the Constitution



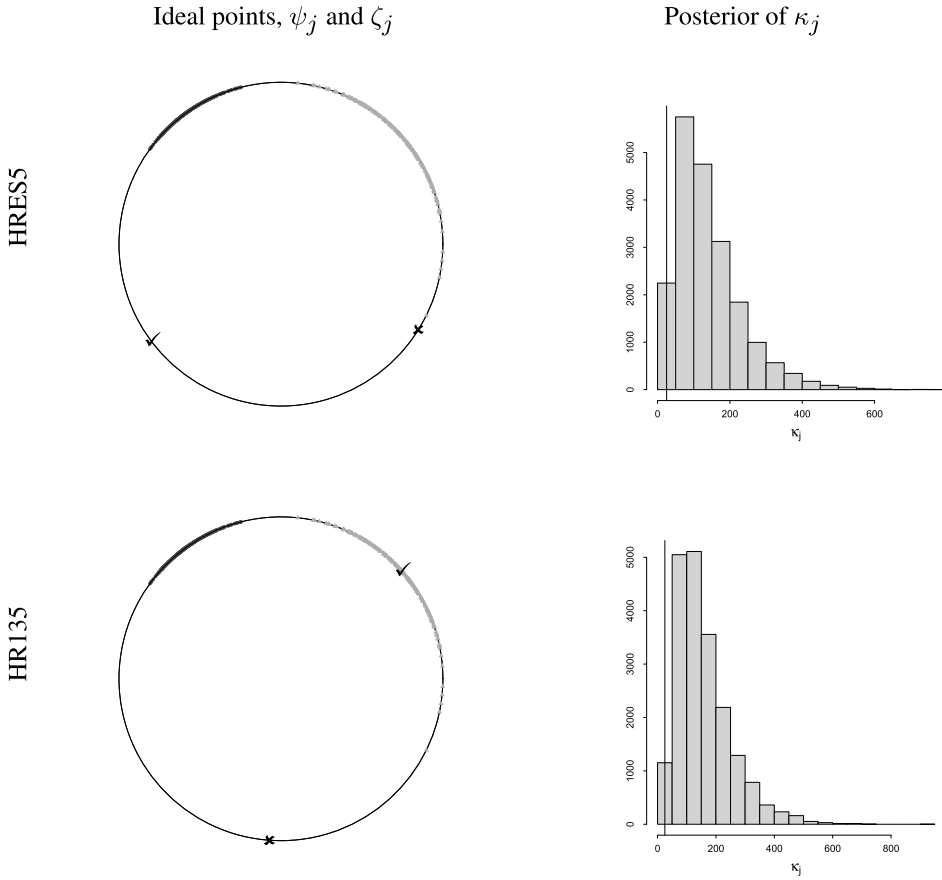


FIG. 5. Two examples of Euclidean voting patterns during the 116th House of Representatives. Graphs on the left column depict the posterior mean ideal point for the legislators (which are the same on both plots), along with the “Yea” and “Nay” positions (represented through a check mark and a cross, respectively). The right columns presents a histogram of samples of the posterior distribution of the corresponding  $\kappa_j$ . The vertical line corresponds to  $1/E(1/\lambda \mid \text{data})$ .

and limited government that has become known as the Tea Party movement (Arceneaux and Nicholson (2012)). Many of these insurgent legislators went on to form congressional member organizations such as the Tea Party Caucus and the House Liberty Caucus (both founded during the 112th Congress, the first in July 2010 and the second in March 2011) as well as the House Freedom Caucus (founded in 2015 during the 114th Congress). These three caucuses are all considered to represent the most extreme wing of the Republican party, and some recent evidence suggests that their members vote like a significantly farther-right third party in Congress (e.g., see Ragusa and Gaspar (2016)). However, as we will see shortly, many of their members are ranked by traditional spatial voting models as mainstream or even centrist Republicans.

Figure 6 compares the rank order of legislators estimated using a one-dimensional model to the rank order estimated by our circular model. For the Democratic Party the ranking generated by both models are reasonably similar. The main outlier is Dennis Kucinich (OH-10), who is ranked as much more liberal by the circular model: his posterior median rank is 2 under the circular model with a 95% credible interval of (1, 6), but it is 76 under the Euclidean model with a 95% credible interval (62, 97). This more extreme ranking fits better with the widely-held perception that Kucinich was one of the most liberal members of the U.S. House of Representatives during this period. In contrast, on the Republican side we see some

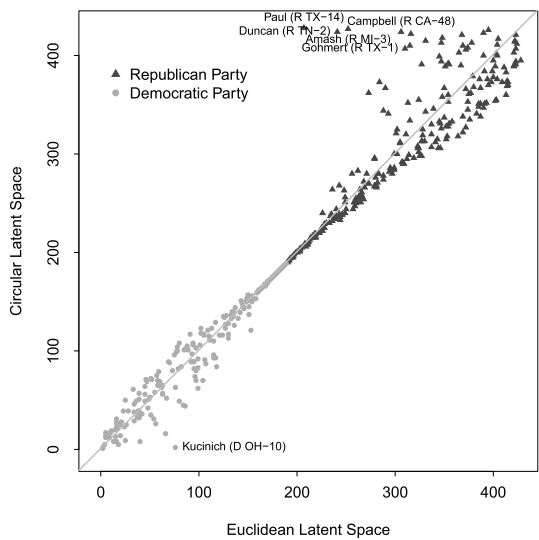


FIG. 6. *Posterior median of the rank-order in the 112th U.S. House of Representatives under the one-dimensional Euclidean (horizontal axis) and the circular (vertical axis) models.*

very large differences between the rankings generated by both models. In particular, we see a large group of legislators that are ranked as much more conservative by the circular model. Figure 7 provides additional details for the 15 Republican legislators for whom the difference in posterior median rankings between the one-dimensional Euclidean and the circular models is largest. It is interesting that, in all cases, the ranks assigned by the circular model are more extreme and that 12 out of the 15 legislators in this list either were members of the Tea Party or the Liberty Caucuses during this Congress or later joined the Freedom Caucus when it

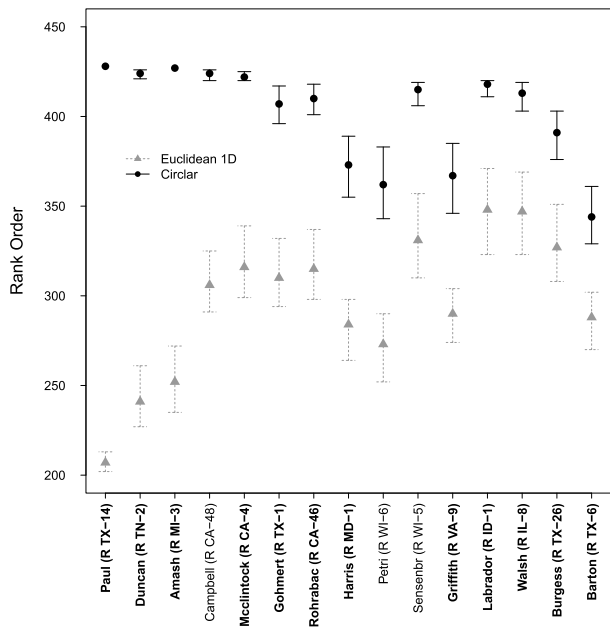


FIG. 7. *Posterior median ranks and associated 95% credible intervals for the 15 Republican legislators in the 112th House for whom the difference in posterior median rankings between the one-dimensional Euclidean and the circular models is largest. Bolded names indicate that the legislator was a member of either the Liberty Caucus, the Freedom Caucus or the Tea Party Caucus at some of point of their career.*

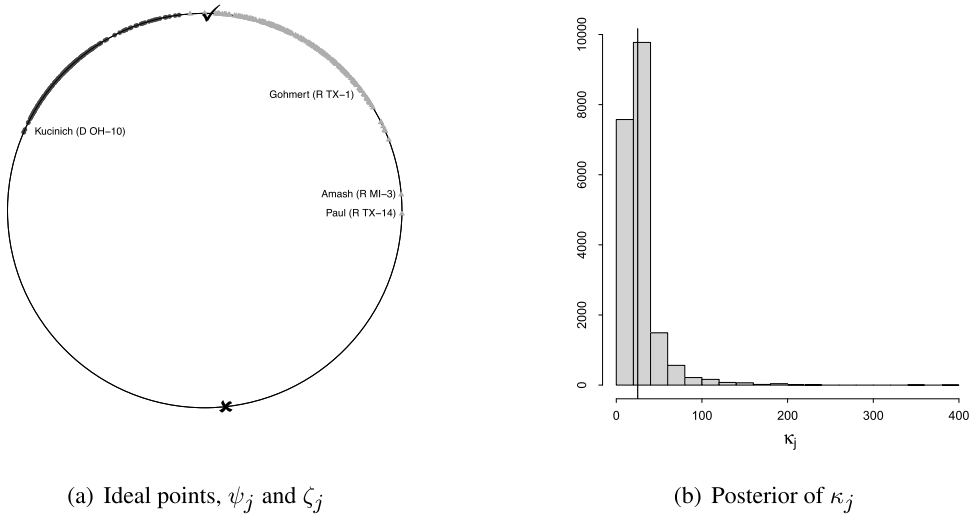


FIG. 8. HR915, one example of circular voting during the 112th House of Representatives. Panel (a) depicts the posterior mean ideal point for the legislators, along with the “Yea” and “Nay” positions (represented through a check mark and a cross, respectively). The names in the graph correspond to the legislators that voted against the measure. Panel (b) presents a histogram of samples of the posterior distribution of the corresponding  $\kappa_j$ . The vertical line corresponds to  $1/E(1/\lambda | \text{data})$ .

was formed. In particular, we must highlight that classifying Ron Paul, Justin Amash (who we already discussed in the previous section) or Jimmy Duncan as centrist (which is the implication from their rank under the Euclidean model) would be very hard to justify based on their stated political positions.

Finally, Figure 8 presents one example of circular voting in the 112th House of Representatives. HR915 is the Jaime Zapata Border Enforcement Security Task Force Act which amended the Homeland Security Act of 2002 to establish the Border Enforcement Security Task Force (BEST). It aimed at facilitating collaboration among federal, state, local, tribal and foreign law enforcement agencies to execute coordinated activities in furtherance of border and homeland security, as well as to enhance information-sharing, including the dissemination of homeland security information among such agencies. Note that HR915 was opposed by representative Kucinich (according to our model, the most extreme Democrat in the House at the time) as well as by both Amash and Massie (in turn, the most extreme Republicans in the House according to our model) and by representative Louie Gohmert (whose ranking also significantly shifts under the circular model). As in Figure 4, the “Nay” position is roughly located opposite to the (circular) average of all ideal points, the “Yea” position is located close to the (circular) average of the ideal points of the legislators that voted in favor of the measure and the posterior distribution of  $\kappa_j$  indicates moderate to low concentration values for the link function.

6.3. *A longitudinal analysis of the contemporary U.S. House of Representatives.* The previous two sections presented two very recent examples of circular voting behavior in the U.S. House of Representatives. We are interested now in understanding how pervasive this behavior has been in modern history. As we mentioned in Section 4.4, the (circular) variance of the ideal points

$$\chi_0 = 1 - \left\{ \left( \frac{1}{I} \sum_{i=1}^I \cos \beta_i \right)^2 + \left( \frac{1}{I} \sum_{i=1}^I \sin \beta_i \right)^2 \right\}^{\frac{1}{2}}$$

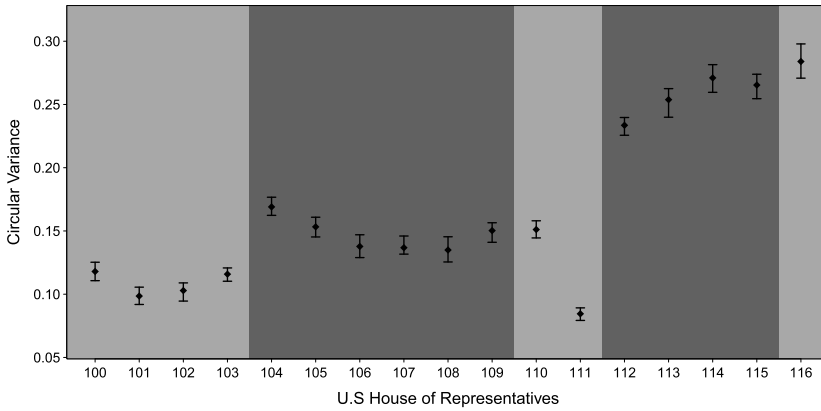


FIG. 9. *Circularity  $\chi_0$  for the 100th to the 116th U.S. House of Representatives. Light gray background indicates a Democratic majority, while dark gray indicates a Republican majority. Results for the 116th House include only the first session.*

provides a natural metric to measure the “circularity” of voting on a given Congress. It is important to note that this metric is useful in this context because, unlike the analogous metric for Euclidean models, it is comparable across Congresses (recall that the utility functions that underlie our formulation are not invariant to rescalings of the policy space).

Figure 9 shows the posterior mean and 95% credible intervals for  $\chi_0$  between the 100th House (1987–1988) and (the first session of) the 116th House (2019). We start the analysis from the mid 1980s because this period sits comfortably after the reforms of the mid-1970s (which included the introduction of electronic voting, leading to a dramatic increase in the number and nature of roll call votes recorded in the chamber). Note that  $\chi_0$  was relatively stable in the late 1980s and early 1990s but then jumped when the control of the chamber switched from Democrats to Republicans with the election of the 104th Congress. It then remained more or less stable during the later half of the 1990s and the 2000s, to then fall during the 111th House and then jump up again to historically high levels from the 112th House on. While the overall increasing trend on  $\chi_0$  agrees with well-known patterns of increasing polarization in Congress, these results also suggest that such an increase in polarization has been accompanied by an increase in the frequency of “extremes voting together,” a phenomenon that has not yet been fully documented or explored in the modern U.S. Congress.

In order to better understand how circular voting has affected each party, we present in Figure 10 the circular variance associated with the ideal of points of both Democrat and Republican legislators,  $\chi_D$  and  $\chi_R$ , respectively. While some of the fluctuations in these metrics roughly match those we observed in Figure 9, it is clear that the overall trend has been for  $\chi_D$  to decrease and for  $\chi_R$  to increase over time. The divergence is particularly stark after the 112th, with the Republican party showing an all-time-high level of intraparty disagreement. Put another way, these results suggest that the Republican party has steadily become more fractious while the Democratic party has tended to unify, particularly over the last 10 years. This pattern is consistent with well-known political processes. On one hand, the steady decrease in  $\chi_D$  between the 100th and the 107th Houses might be seen as the upshot of the migration of the remaining former Southern Democrats to the Republican party. On the other hand, the large increase in  $\chi_R$  starting in the 112th can be understood as a consequence of the rise of the Tea Party movement (recall our discussion in Section 6.2).

To conclude this section, we present in Table 4 the value of the Watanabe–Akaike information criteria (WAIC, see Gelman, Hwang and Vehtari (2014), Watanabe (2010), Watanabe (2013)) associated with both the circular and one-dimensional Euclidean models fitted to the

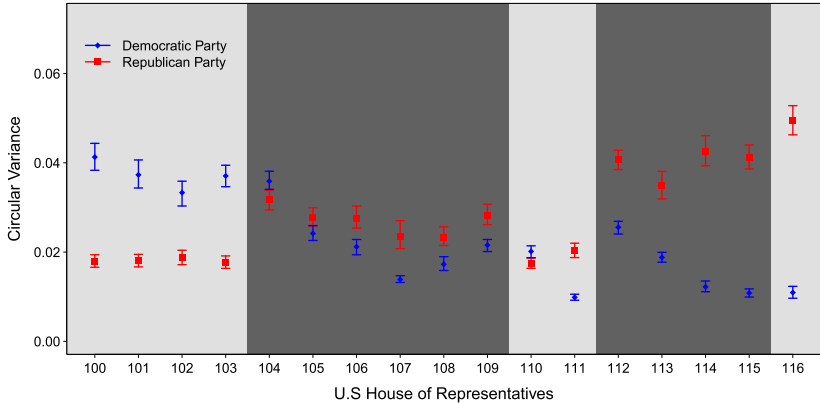


FIG. 10. Within-party circular variances,  $\chi_D$  and  $\chi_R$ , for the 100th to the 116th U.S. House of Representatives. Light grey background indicates a Democratic majority, while dark grey indicates a Republican majority. Results for the 116th House include only the first session.

data from the 100th to the 116th U.S. House of Representatives. Similarly to the well-known Akaike information criteria (AIC) and the Bayesian information criteria (BIC), WAIC balances goodness of fit against model complexity. However, unlike the AIC and the BIC, the WAIC is well suited for hierarchical models where the number of effective parameters can be much smaller than the headline number. Compared to the deviance information criteria (DIC), the WAIC has the advantage of being invariant to reparameterizations of the model (Spiegelhalter et al. (2014), Gelman, Hwang and Vehtari (2014)). We employ here the so-called “grouped” WAIC in which the the vector of all votes for a given legislator are treated as one observation (e.g., see Gelman, Hwang and Vehtari ((2014), page 1009)).

The results in Table 4 are extremely interesting and further reinforce the observation that there has been a fundamental change in the American political environment since 2010. In-

TABLE 4  
*Grouped WAIC for the the circular and Euclidean model  
fitted to the 100th to 116th U.S. House of Representatives.  
Results for the 116th House include only the first session*

| House | I   | J    | Euclidean (1D)  | Circular        |
|-------|-----|------|-----------------|-----------------|
| 100th | 425 | 939  | <b>−94,801</b>  | −98,634         |
| 101th | 429 | 879  | <b>−103,745</b> | −106,598        |
| 102th | 431 | 901  | <b>−103,952</b> | −108,156        |
| 103th | 430 | 1094 | <b>−114,457</b> | −117,125        |
| 104th | 428 | 1321 | <b>−132,138</b> | −135,108        |
| 105th | 426 | 1166 | <b>−114,196</b> | −115,590        |
| 106th | 432 | 1209 | <b>−109,537</b> | −110,554        |
| 107th | 428 | 990  | −71,230         | <b>−70,242</b>  |
| 108th | 430 | 1218 | <b>−73,137</b>  | −73,143         |
| 109th | 430 | 1210 | −86,808         | <b>−86,721</b>  |
| 110th | 424 | 1865 | <b>−97,252</b>  | −98,775         |
| 111th | 426 | 1647 | <b>−74,435</b>  | −74,976         |
| 112th | 428 | 1602 | −120,633        | <b>−120,313</b> |
| 113th | 424 | 1202 | −75,718         | <b>−75,249</b>  |
| 114th | 431 | 1322 | −68,118         | <b>−67,198</b>  |
| 115th | 450 | 1207 | −56,104         | <b>−54,994</b>  |
| 116th | 432 | 700  | −33,131         | <b>−31,403</b>  |

deed, note that between the 100th and the 111th Houses, the Euclidean model tends to dominate our circular model, and, in the cases in which that is not the case (the 107th and 109th Houses), the advantage of the circular model is relatively modest. On the other hand, since the 112th House, the circular model is clearly preferred by the data. This roughly coincides with the pattern observed for  $\chi_0$  in Figure 9, suggesting that the circular variance of the ideal points is indeed a good metric of “circularity” in voting patterns.

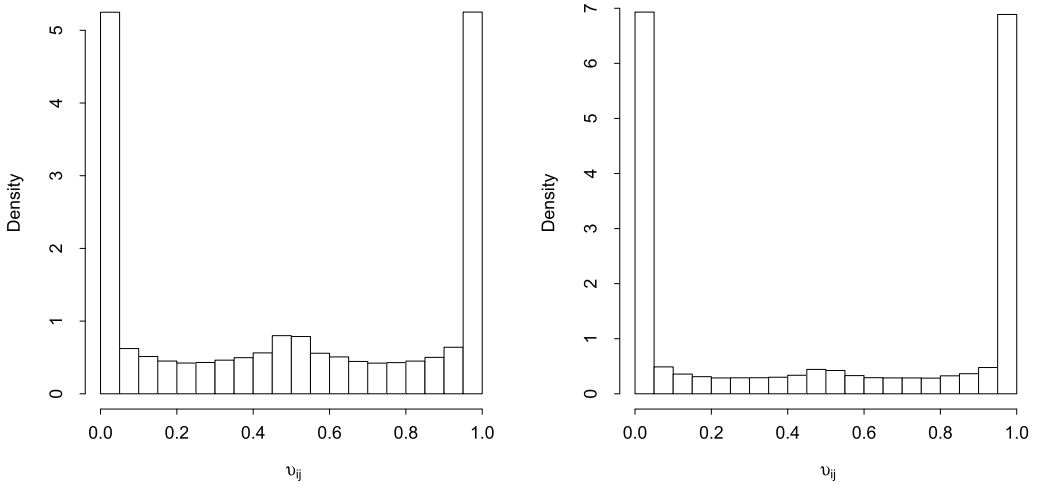
**6.4. Sensitivity analysis.** In order to understand the effect of the priors on our results, we conducted a sensitivity analysis by refitting the model under three alternative priors for each of the 112th and 116th Houses. First, we consider a Gamma prior with shape parameter 7 and scale parameter 1 for  $\omega_\beta$  (so that  $\Pr(-\pi/2 < \beta_i < \pi/2) \approx 0.995$  a priori) and set the mean of  $1/\lambda$  to  $1/100$ . Relatively speaking, this prior favors configurations that are closer to those generated by the one-dimensional Euclidean model. Second, we consider an exponential prior with mean 2 for  $\omega_\beta$  (so that  $\Pr(-\pi/2 < \beta_i < \pi/2) \approx 0.80$  a priori) and set the mean of  $1/\lambda$  to  $1/25$ . Compared to the first hyperprior, this second one favors circular configurations. Finally, we consider a combination of an exponential prior with mean 10 for  $\omega_\beta$  (just as in our original prior) and a second exponential prior with mean 2 for  $1/\lambda$ .

Figure 11 presents histograms of 100,000 samples simulated from the (marginal) prior distribution on  $v_{i,j} = G_{\kappa_j}(e_{i,j}(\psi_j, \zeta_j, \beta_i))$  (the probability of a positive vote) induced by these three alternative hyperparameter combinations as well as our original set of hyperparameters. These marginal distributions (which, a priori, are identical for every  $i$  and  $j$ ) are a useful tool to understand how the priors on the latent space impact the model. In this case, note that the first two alternative sets of hyperparameters lead to priors on  $v_{i,j}$  that look look similar to our original one. These priors place high probability on values close to 0 and 1, a behavior that would be expected in roll votes, especially in legislative bodies with a high level of polarization, such as the U.S. House of Representatives (see, e.g., [Spirling and Quinn \(2010\)](#)). On the other hand, our third alternative set of hyperparameters yields a very different prior for  $v_{i,j}$ , one that places most of its mass on values close to  $1/2$ . In spite of these differences, posterior inferences seem to be robust across all four priors we considered. Detailed results can be seen in Section 2 of the Supplementary Materials ([Yu and Rodriguez \(2021a\)](#)).

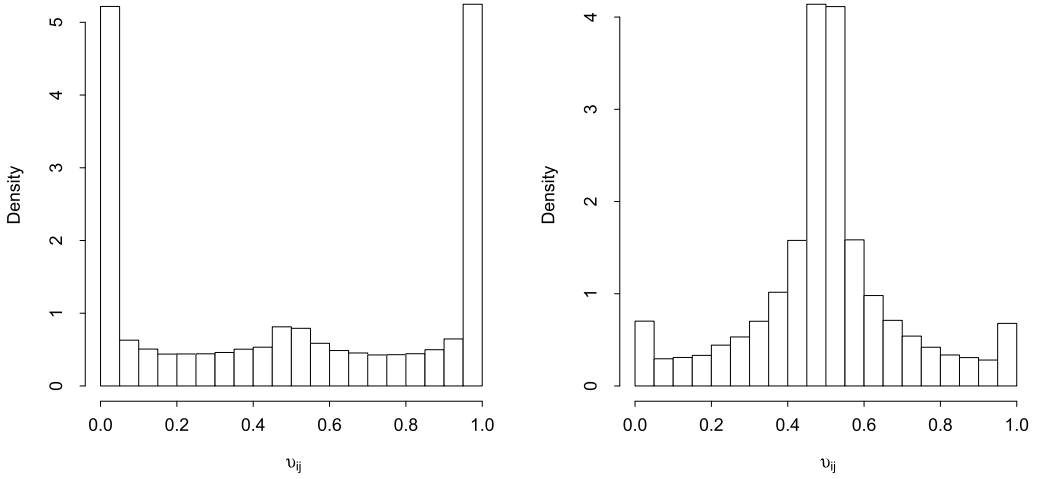
**7. Discussion.** Our results suggest that the circular voting model developed in this paper provides a better explanation for voting patterns in the modern U.S. House of Representatives than traditional one-dimensional voting models. In particular, we find a pattern of increasing circularity driven by the raise of extreme ideological factions willing to vote against the mainstream of the party (especially among Republicans) which seems to have gone hand in hand with increasing polarization in the chamber.

One referee astutely wondered whether there were any instance of the circular model misplacing a legislator by assigning a moderate a more extreme position than the Euclidean model. Across the two Houses that we studied in detail (the 112th and 116th), the only legislator for whom this appeared to be the case is representative John Campbell (R CA-48) in the 112th (see Figure 7). Most analysts would not consider Campbell a Republican firebrand. Therefore, having him appear in our list of largest changes was a bit of a surprise. However, representative Campbell did have a history of voting with (sometimes extreme) liberals on some issues. Relevant examples include being one of 15 Republican House members to vote in favor of repealing the United States military’s “Don’t Ask, Don’t Tell” ban on openly gay service members on December 2010 and voting against the FY12 National Defense Authorization Act in protest against a controversial provision allowing the military to indefinitely detain American citizens and others without trial. In fact, the placement of Campbell under the circular model serves to illustrate how our model works. Unlike the Euclidean model, our





(a) Original:  $\omega_\beta \sim \text{Gam}(1, 0.1)$ ,  $1/\lambda \sim \text{Gam}(1, 25)$  (b) Alternative 1:  $\omega_\beta \sim \text{Gam}(7, 1)$ ,  $1/\lambda \sim \text{Gam}(1, 100)$



(c) Alternative 2:  $\omega_\beta \sim \text{Gam}(1, 0.5)$ ,  $1/\lambda \sim \text{Gam}(1, 25)$  (d) Alternative 3:  $\omega_\beta \sim \text{Gam}(1, 0.1)$ ,  $1/\lambda \sim \text{Gam}(1, 0.5)$

FIG. 11. Marginal prior distribution on  $v_{i,j} = G_{\kappa_j}(e_{i,j}(\psi_j, \zeta_j, \beta_i))$  (the probability of a positive vote) induced by our original prior distribution as well as those induced by three alternative hyperparameter values used in our sensitivity analysis.

circular model treats legislators that vote with centrists of the other party differently from how it treats those that vote with their extreme partisans. In the former case, the relative position of the legislator can be expected to be very similar under both models. In the latter, the position of the legislator can be expected to be more extreme under the circular model. Hence, what our results reveal about representative Campbell is that, in spite of whatever policy positions he publicly expressed, his voting record is nonetheless consistent with sporadic votes in coalition with liberal Democrats.

In our interactions with various colleagues, one criticism that has been leveled at times is whether the circular assumption makes mechanistic sense *ex ante* and, in particular, whether the Horseshoe Theory can be applied in the context of the U.S. House of Representatives.

We view this criticism as a reflection of the type deductive thinking that tends to dominate in political science. In contrast, this paper posits an inductive approach to the problem in which we use data to evaluate the empirical support for the Horseshoe Theory. The main contribution of this paper is to provide a solid statistical methodology that enables this kind empirical explorations, under the much laxer assumption that a circular voting space is at least minimally plausible. We recognize that taking an inductive approach begs the question of whether there are alternative explanations (e.g., heavy-tailed random shocks to the latent utilities or nonmonotonic utilities functions) for the “extremes voting together” phenomenon. We agree that further exploration of these questions is key but also see such endeavors as beyond the scope of this particular paper which is focused on introducing the basic methodology needed to fit and assess evidence related to models that rely on circular policy spaces. Furthermore, the fact that our results partially challenge the ex-ante opinion that the Horseshoe Theory might not be a good fit in the U.S. that some experts might hold should be seen as a finding worth further investigation rather than a reason to discard the underlying methodology enabling those conclusion.

Because our motivation comes from the study of the U.S. Congress, our focus has been solely on investigating the impact of geometry on intrinsically unidimensional models. However, for other applications, learning the dimensionality of the latent spherical space and understanding the trade-offs between dimensionality and geometry might be critical tasks. Addressing these questions is, however, nontrivial. In particular, priors on the latent space need to be constructed very carefully, particularly if dimension selection is important. This is part of our current work in progress.

The Markov chain Monte Carlo algorithm devised in Section 5 seems to perform quite well in the examples we considered. However, there are a number of possible improvement, from implementing joint HMC steps for blocks of parameters rather than for individual ones, to adapting the ideas behind the No-U-Turn sampler of [Hoffman and Gelman \(2014\)](#) to the context GHMC algorithms in order to minimize the amount of tuning required. We will explore some of these alternative computational approaches in the future as we improve our software.

Our model makes no attempt to explicitly model how or whether the voting patterns of individual legislators have evolved over time, neither within a given House, nor across Houses. In the case of Euclidean models, such questions have been addressed before for (e.g., see [Martin and Quinn \(2002\)](#) and [Lofland, Rodríguez and Moser \(2017\)](#)). We will explore similar extensions of our circular models in our future work.

## APPENDIX: HAUSDORFF MEASURES AND THEIR GRADIENTS

The density of the full conditional distribution for  $\beta_i$  is given by

$$\begin{aligned} p(\beta_i | \dots) &\propto \exp\{\omega_\beta \beta_i\} \\ &\times \prod_{j=1}^J [G_{\kappa_j}(\{\arccos(\cos(\zeta_j - \beta_i))\}^2 - \{\arccos(\cos(\psi_j - \beta_i))\}^2)]^{y_{i,j}} \\ &\times [1 - G_{\kappa_j}(\{\arccos(\cos(\zeta_j - \beta_i))\}^2 - \{\arccos(\cos(\psi_j - \beta_i))\}^2)]^{1-y_{i,j}}. \end{aligned}$$

Then, the density of the associated Hausdorff measure in  $\mathbb{R}^2$  is given by

$$\begin{aligned} p(\mathbf{x}_{\beta_i} | \dots) &\propto \exp\{\boldsymbol{\eta}_\beta^T \mathbf{x}_{\beta_i}\} \prod_{j=1}^J [G_{\kappa_j}(\{\arccos(\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})\}^2 - \{\arccos(\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})\}^2)]^{y_{i,j}} \\ &\times [1 - G_{\kappa_j}(\{\arccos(\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})\}^2 - \{\arccos(\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})\}^2)]^{1-y_{i,j}}, \quad \mathbf{x}_{\beta_i}^T \mathbf{x}_{\beta_i} = 1, \end{aligned}$$

where  $\boldsymbol{\eta}_\beta^T = (\omega_\beta, 0)$ ,  $\mathbf{z}_{\psi_j}^T = (\cos \psi_j, \sin \psi_j)$ ,  $\mathbf{z}_{\zeta_j}^T = (\cos \zeta_j, \sin \zeta_j)$  and the mapping between  $\beta_i$  and  $\mathbf{x}_{\beta_i}$  is given by  $\mathbf{x}_{\beta_i}^T = (\cos \beta_i, \sin \beta_i)$ . Hence, the gradient of the Hausdorff measure is simply

$$\nabla \log_{\mathcal{H}} p(\mathbf{x}_{\beta_i} | \cdots) = \boldsymbol{\eta}_\beta + \sum_{j=1}^J \begin{pmatrix} y_{i,j} e'_{i,j,1} \frac{g_{\kappa_j}(e_{i,j})}{G_{\kappa_j}(e_{i,j})} - (1 - y_{i,j}) e'_{i,j,1} \frac{g_{\kappa_j}(e_{i,j})}{1 - G_{\kappa_j}(e_{i,j})} \\ y_{i,j} e'_{i,j,2} \frac{g_{\kappa_j}(e_{i,j})}{G_{\kappa_j}(e_{i,j})} - (1 - y_{i,j}) e'_{i,j,2} \frac{g_{\kappa_j}(e_{i,j})}{1 - G_{\kappa_j}(e_{i,j})} \end{pmatrix},$$

where

$$e_{i,j} = \{\arccos(\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})\}^2 - \{\arccos(\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})\}^2, \\ e'_{i,j} = \begin{pmatrix} e'_{i,j,1} \\ e'_{i,j,2} \end{pmatrix} = \frac{2 \arccos(\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})}{\sqrt{1 - (\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})^2}} \mathbf{z}_{\psi_j} - \frac{2 \arccos(\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})}{\sqrt{1 - (\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})^2}} \mathbf{z}_{\zeta_j}.$$

Next, the density of the full conditional distribution for  $\mathbf{z}_{\zeta_j}$  in terms of Hausdorff measure in  $\mathbb{R}^2$  is given by

$$p(\mathbf{z}_{\zeta_j} | \cdots) \propto \prod_{i=1}^I G_{\kappa_j}(\{\arccos(\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})\}^2 - \{\arccos(\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})\}^2) \\ \times [1 - G_{\kappa_j}(\{\arccos(\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})\}^2 - \{\arccos(\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})\}^2)]^{1-y_{i,j}}, \quad \mathbf{z}_{\zeta_j}^T \mathbf{z}_{\zeta_j} = 1,$$

and the gradient of the Hausdorff measure is

$$\nabla \log_{\mathcal{H}} p(\mathbf{z}_{\zeta_j} | \cdots) = \sum_{i=1}^I \begin{pmatrix} y_{i,j} \delta'_{i,j,1} \frac{g_{\kappa_j}(\delta_{i,j})}{G_{\kappa_j}(\delta_{i,j})} - (1 - y_{i,j}) \delta'_{i,j,1} \frac{g_{\kappa_j}(\delta_{i,j})}{1 - G_{\kappa_j}(\delta_{i,j})} \\ y_{i,j} \delta'_{i,j,2} \frac{g_{\kappa_j}(\delta_{i,j})}{G_{\kappa_j}(\delta_{i,j})} - (1 - y_{i,j}) \delta'_{i,j,2} \frac{g_{\kappa_j}(\delta_{i,j})}{1 - G_{\kappa_j}(\delta_{i,j})} \end{pmatrix},$$

where  $\delta_{i,j} = \{\arccos(\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})\}^2 - \{\arccos(\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})\}^2$ ,

$$\delta'_{i,j} = \begin{pmatrix} \delta'_{i,j,1} \\ \delta'_{i,j,2} \end{pmatrix} = -\frac{2 \arccos(\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})}{\sqrt{1 - (\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})^2}} \mathbf{x}_{\beta_i}.$$

Lastly, the density of the full conditional distribution for  $\mathbf{z}_{\psi_j}$  in terms of Hausdorff measure in  $\mathbb{R}^2$  is given by

$$p(\mathbf{z}_{\psi_j} | \cdots) \propto \prod_{i=1}^I G_{\kappa_j}(\{\arccos(\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})\}^2 - \{\arccos(\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})\}^2) \\ \times [1 - G_{\kappa_j}(\{\arccos(\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})\}^2 - \{\arccos(\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})\}^2)]^{1-y_{i,j}}, \quad \mathbf{z}_{\psi_j}^T \mathbf{z}_{\psi_j} = 1,$$

and the gradient of the Hausdorff measure is

$$\nabla \log_{\mathcal{H}} p(\mathbf{z}_{\psi_j} | \cdots) = \sum_{i=1}^I \begin{pmatrix} y_{i,j} \gamma'_{i,j,1} \frac{g_{\kappa_j}(\gamma_{i,j})}{G_{\kappa_j}(\gamma_{i,j})} - (1 - y_{i,j}) \gamma'_{i,j,1} \frac{g_{\kappa_j}(\gamma_{i,j})}{1 - G_{\kappa_j}(\gamma_{i,j})} \\ y_{i,j} \gamma'_{i,j,2} \frac{g_{\kappa_j}(\gamma_{i,j})}{G_{\kappa_j}(\gamma_{i,j})} - (1 - y_{i,j}) \gamma'_{i,j,2} \frac{g_{\kappa_j}(\gamma_{i,j})}{1 - G_{\kappa_j}(\gamma_{i,j})} \end{pmatrix},$$

where

$$\gamma_{i,j} = \{\arccos(\mathbf{z}_{\zeta_j}^T \mathbf{x}_{\beta_i})\}^2 - \{\arccos(\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})\}^2,$$

$$\gamma'_{i,j} = \begin{pmatrix} \gamma'_{i,j,1} \\ \gamma'_{i,j,2} \end{pmatrix} = \frac{2 \arccos(\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})}{\sqrt{1 - (\mathbf{z}_{\psi_j}^T \mathbf{x}_{\beta_i})^2}} \mathbf{x}_{\beta_i}.$$

**Acknowledgments.** Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

**Funding.** AR was partially supported by NSF Grant DMS-2114727, DMS-2114729 and DMS-2023495.

## SUPPLEMENTARY MATERIAL

**Supplementary Material A** (DOI: [10.1214/21-AOAS1454SUPPA](https://doi.org/10.1214/21-AOAS1454SUPPA); .pdf). We discuss our convergence, mixing analyses and sensitivity analysis for the MCMC runs on which the results in Sections 6.1 and 6.2 are based. In addition, we report the outcomes of two simulation studies designed to illustrate that it is possible to differentiate between a circular and a 2D Euclidean latent space, and to illustrate the connection between circular and higher-dimensional Euclidean models.

**Supplementary Material B** (DOI: [10.1214/21-AOAS1454SUPPB](https://doi.org/10.1214/21-AOAS1454SUPPB); .zip). R and Rcpp source code for the models implemented in this paper and data obtained from <https://voteview.com/>.

## REFERENCES

- ARCENEUX, K. and NICHOLSON, S. P. (2012). Who wants to have a tea party? The who, what, and why of the Tea Party movement. *PS Polit. Sci. Polit.* **45** 700–710.
- BAFUMI, J., GELMAN, A., PARK, D. K. and KAPLAN, N. (2005). Practical issues in implementing and understanding Bayesian ideal point estimation. *Polit. Anal.* **13** 171–87.
- BELKIN, M. and NIYOGI, P. (2002). Laplacian eigenmaps and spectral techniques for embedding and clustering. In *Advances in Neural Information Processing Systems* 585–591.
- BESKOS, A., PILLAI, N., ROBERTS, G., SANZ-SERNA, J.-M. and STUART, A. (2013). Optimal tuning of the hybrid Monte Carlo algorithm. *Bernoulli* **19** 1501–1534. [MR3129023 https://doi.org/10.3150/12-BEJ414](https://doi.org/10.3150/12-BEJ414)
- BETANCOURT, M. J., BYRNE, S. and GIROLAMI, M. (2014). Optimizing the integrator step size for Hamiltonian Monte Carlo. Preprint. Available at [arXiv:1411.6669](https://arxiv.org/abs/1411.6669).
- BYRNE, S. and GIROLAMI, M. (2013). Geodesic Monte Carlo on embedded manifolds. *Scand. J. Stat.* **40** 825–845. [MR3145120 https://doi.org/10.1111/sjos.12036](https://doi.org/10.1111/sjos.12036)
- CARROLL, R., LEWIS, J. B., LO, J., POOLE, K. T. and ROSENTHAL, H. (2013). The structure of utility in spatial models of voting. *Amer. J. Polit. Sci.* **57** 1008–1028.
- CLINTON, J. D. and JACKMAN, S. (2009). To simulate or NOMINATE? *Legis. Stud. Q.* **34** 593–621.
- CLINTON, J., JACKMAN, S. and RIVERS, D. (2004). The statistical analysis of roll call data. *Am. Polit. Sci. Rev.* **98** 355–370.
- CRANE, H. (2017). A hidden Markov model for latent temporal clustering with application to ideological alignment in the U.S. Supreme Court. *Comput. Statist. Data Anal.* **110** 19–36. [MR3612605 https://doi.org/10.1016/j.csda.2016.12.010](https://doi.org/10.1016/j.csda.2016.12.010)
- DAVIS, O. A., HINICH, M. J. and ORDESHOOK, P. C. (1970). An expository development of a mathematical model of the electoral process. *Am. Polit. Sci. Rev.* **64** 426–448.
- DUCK-MAYR, J. and MONTGOMERY, J. M. (2020). Ends against the middle: Scaling votes when ideological opposites behave the same for antithetical reasons. Technical report, Department of Political Science, Washington Univ. in St. Louis. Available at <https://polmeth.org/publications/ends-against-middle-scaling-votes-when-ideological-opposites-behave-same>.

- EGUIA, J. X. (2013). Challenges to the standard Euclidean spatial model. In *Advances in Political Economy* 169–180. Springer, Berlin.
- ENELOW, J. M. and HINICH, M. J. (1984). *The Spatial Theory of Voting: An Introduction*. CUP Archive, Cambridge.
- FOX, J.-P. (2010). *Bayesian Item Response Modeling: Theory and Applications. Statistics for Social and Behavioral Sciences*. Springer, New York. MR2657265 <https://doi.org/10.1007/978-1-4419-0742-4>
- GELMAN, A., HWANG, J. and VEHTARI, A. (2014). Understanding predictive information criteria for Bayesian models. *Stat. Comput.* **24** 997–1016. MR3253850 <https://doi.org/10.1007/s11222-013-9416-2>
- GELMAN, A. and RUBIN, D. B. (1992). Inferences from iterative simulation using multiple sequences. *Statist. Sci.* **7** 457–472.
- GELMAN, A., CARLIN, J. B., STERN, H. S., DUNSON, D. B., VEHTARI, A. and RUBIN, D. B. (2014). *Bayesian Data Analysis*, 3rd ed. *Texts in Statistical Science Series*. CRC Press, Boca Raton, FL. MR3235677
- GEWEKE, J. F. and SINGLETON, K. J. (1981). Maximum likelihood “confirmatory” factor analysis of economic time series. *Internat. Econom. Rev.* **22** 37–54. MR0614346 <https://doi.org/10.2307/2526134>
- GUIMERÀ, R. and SALES-PARDO, M. (2011). Justice blocks and predictability of US Supreme Court votes. *PLoS ONE* **6** e27188.
- HARE, C. and POOLE, K. T. (2014). The polarization of contemporary American politics. *Polity* **46** 411–429.
- HECKMAN, J. J. and SNYDER JR., J. M. (1996). Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. Technical report, National bureau of economic research.
- HOFFMAN, M. D. and GELMAN, A. (2014). The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **15** 1593–1623. MR3214779
- HUMPHREYS, M. and LAVER, M. (2010). Spatial models, cognitive metrics, and majority rule equilibria. *Br. J. Polit. Sci.* **40** 11–30.
- JACKMAN, S. (2001). Multidimensional analysis of roll call data via Bayesian simulation: Identification, estimation, inference, and model checking. *Polit. Anal.* **9** 227–241.
- JESSEE, S. A. (2012). *Ideology and Spatial Voting in American Elections*. Cambridge Univ. Press, Cambridge.
- KARPOWITZ, C. F., MONSON, J. Q., PATTERSON, K. D. and POPE, J. C. (2011). Tea time in America? The impact of the Tea Party movement on the 2010 midterm elections. *PS Polit. Sci. Polit.* **44** 303–309.
- KINGMA, D. P., WELLING, M. et al. (2019). An introduction to variational autoencoders. *Found. Trends Mach. Learn.* **12** 307–392.
- KRAMER, M. A. (1991). Nonlinear principal component analysis using autoassociative neural networks. *AIChE J.* **37** 233–243.
- LAWRENCE, N. (2005). Probabilistic non-linear principal component analysis with Gaussian process latent variable models. *J. Mach. Learn. Res.* **6** 1783–1816. MR2249872
- LEWIS, J. (2019a). Why are Ocasio-Cortez, Omar, Pressley, and Talib estimated to be moderates by NOMINATE? Available at [https://voteview.com/articles/Ocasio-Cortez\\_Omar\\_Pressley\\_Tlaib](https://voteview.com/articles/Ocasio-Cortez_Omar_Pressley_Tlaib).
- LEWIS, J. (2019b). Why is Alexandria Ocasio-Cortez estimated to be a moderate by NOMINATE? Available at [https://voteview.com/articles/ocasio\\_cortez](https://voteview.com/articles/ocasio_cortez).
- LOFLAND, C. L., RODRÍGUEZ, A. and MOSER, S. (2017). Assessing differences in legislators’ revealed preferences: A case study on the 107th U.S. Senate. *Ann. Appl. Stat.* **11** 456–479. MR3634331 <https://doi.org/10.1214/16-AOAS951>
- MARTIN, A. D. and QUINN, K. M. (2002). Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999. *Polit. Anal.* **10** 134–153.
- MCCARTY, N., POOLE, K. T. and ROSENTHAL, H. (2016). *Polarized America: The Dance of Ideology and Unequal Riches*, 2nd ed. MIT Press, Cambridge.
- MCCORMICK, T. H. and ZHENG, T. (2015). Latent surface models for networks using aggregated relational data. *J. Amer. Statist. Assoc.* **110** 1684–1695. MR3449064 <https://doi.org/10.1080/01621459.2014.991395>
- MCFADDEN, D. (1973). Conditional logit analysis of qualitative choice behavior. In *Frontiers of Economics* (P. Zarembka, ed.) 105–142. Institute of Urban and Regional Development, Univ. California.
- MOSER, S., RODRIGUEZ, A. and LOFLAND, C. L. (2021). Multiple ideal points: Revealed preferences in different domains. *Polit. Anal.* **29** 139–166.
- PIERRE, F. J. (2002). *Le siècle des idéologies*. Pocket, Paris.
- POOLE, K. T. and ROSENTHAL, H. (1985). A spatial model for legislative roll call analysis. *Amer. J. Polit. Sci.* **29** 357–384.
- POOLE, K. T. and ROSENTHAL, H. (1991). Patterns of congressional voting. *Amer. J. Polit. Sci.* **35** 228–278.
- POOLE, K. T. and ROSENTHAL, H. (1997). *Congress: A Political-Economic History of Roll Call Voting*. Oxford Univ. Press, London.
- RAGUSA, J. M. and GASPAR, A. (2016). Where’s the Tea Party? An examination of the Tea Party’s voting behavior in the House of Representatives. *Polit. Res. Q.* **69** 361–372.

- RODRÍGUEZ, A. and MOSER, S. (2015). Measuring and accounting for strategic abstentions in the US Senate, 1989–2012. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **64** 779–797. [MR3415950](#) <https://doi.org/10.1111/rssc.12099>
- ROWEIS, S. T. and SAUL, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science* **290** 2323–2326.
- SKOCPOL, T. and WILLIAMSON, V. (2016). *The Tea Party and the Remaking of Republican Conservatism*. Oxford Univ. Press, London.
- SMITH, A. L., ASTA, D. M. and CALDER, C. A. (2019). The geometry of continuous latent space models for network data. *Statist. Sci.* **34** 428–453. [MR4017522](#) <https://doi.org/10.1214/19-STS702>
- SPIEGELHALTER, D. J., BEST, N. G., CARLIN, B. P. and VAN DER LINDE, A. (2014). The deviance information criterion: 12 years on. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 485–493. [MR3210727](#) <https://doi.org/10.1111/rssb.12062>
- SPIRLING, A. and MCLEAN, I. (2007). UK OC OK? Interpreting optimal classification scores for the UK House of Commons. *Polit. Anal.* **15** 85–96.
- SPIRLING, A. and QUINN, K. (2010). Identifying intraparty voting blocs in the U.K. House of Commons. *J. Amer. Statist. Assoc.* **105** 447–457. [MR2724838](#) <https://doi.org/10.1198/jasa.2009.ap07115>
- TAYLOR, J. (2006). *Where Did the Party Go?: William Jennings Bryan, Hubert Humphrey, and the Jeffersonian Legacy*. Univ. Missouri Press, Columbia.
- TENENBAUM, J. B., DE SILVA, V. and LANGFORD, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science* **290** 2319–2323.
- VAN DER LINDEN, W. J. and HAMBLETON, R. K., eds. (1997). *Handbook of Modern Item Response Theory*. Springer, New York. [MR1601043](#) <https://doi.org/10.1007/978-1-4757-2691-6>
- WATANABE, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J. Mach. Learn. Res.* **11** 3571–3594. [MR2756194](#)
- WATANABE, S. (2013). A widely applicable Bayesian information criterion. *J. Mach. Learn. Res.* **14** 867–897. [MR3049492](#)
- WEISBERG, H. F. (1974). Dimensionland: An excursion into spaces. *Amer. J. Polit. Sci.* **18** 743–776.
- WILSON, M. and DE BOECK, P. (2004). Descriptive and explanatory item response models. In *Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach*. *Stat. Soc. Sci. Public Policy* 43–74. Springer, New York. [MR2083195](#) [https://doi.org/10.1007/978-1-4757-3990-9\\_2](https://doi.org/10.1007/978-1-4757-3990-9_2)
- YU, X. and RODRIGUEZ, A. (2021a). Supplement to “Spatial voting models in circular spaces: A case study of the U.S. House of Representatives.” <https://doi.org/10.1214/21-AOAS1454SUPPA>
- YU, X. and RODRIGUEZ, A. (2021b). Supplement to source code “Spatial voting models in circular spaces: A case study of the U.S. House of Representatives.” <https://doi.org/10.1214/21-AOAS1454SUPPB>
- ZHANG, Z. and ZHA, H. (2004). Principal manifolds and nonlinear dimensionality reduction via tangent space alignment. *SIAM J. Sci. Comput.* **26** 313–338. [MR2114346](#) <https://doi.org/10.1137/S1064827502419154>