Modeling preferences using legislative voting in the presence of missing data

Abel Rodríguez [*] Scott Moser [†]

24 August 2013

Abstract

Strategic abstentions are a poorly understood feature of legislative voting records. This paper discusses a spatial model for legislator's revealed preferences that accounts for abstentions driven by competing principals. A particularly appealing feature of our model is its ability to identify legislators that consistently engage in strategic abstentions, as well as bills for which the position of the legislator in policy space is a key driver of abstentions. We illustrate the performance of our model through the analysis of two datasets, one from the period leading to the Second Reform Act of 1867 in the UK House of Commons, and one from the second session of the 108^{th} U.S. Senate.

1 Introduction

Statistical methods for the analysis of voting records in legislative and judicial environments have become one of the most important quantitative tools in political science. These methods are used for exploratory purposes to describe the behavior of the legislators *vis-à-vis* their party affiliation (e.g., see Jenkkins, 2006) and to understand how voter's opinions evolve over time (e.g., see Martin

[*] Associate Professor, Department of Applied Mathematics and Statistics, University of California, Santa Cruz, 1156 High Street, Mailstop SOE2, Santa Cruz, CA, 95064; Ph.: (831) 459-1047, Fax: (831) 459-4482; Email: `abel@soe.ucsc.edu.edu` URL:`http://soe.ucsc.edu/~abel`

[†] Assistant Professor, Department of Government, University of Texas at Austin, 1 University Sta. A1800, Austin, TX, 78712; Ph.: (512) 232-7305, Fax: (512) 471-1061; Email: `smoser@austin.utexas.edu` URL:`http://smoser.webhost.utexas.edu/`

& Quinn, 2002 and McCarty et al., 2006), as well as in confirmatory settings to test specific theories of legislative or judicial behavior (e.g., see Schickler, 2000 and Wright & Schaffner, 2002).

The so-called "spatial models" of voter behavior (Davis et al., 1970; Enelow & Hinich, 1984) form the basis for the most popular class of statistical models used for the analysis of voting records. In spatial models, legislators and bills are represented as (unknown) points in a latent low-dimensional space.[1] The goal of the statistical model is to recover the location of these points, with a particular focus on the location of the legislators' "ideal points" in a Euclidean "policy space." Examples of spatial voting models include the Heckman-Snyder linear estimator (Heckerman & Snyder, 1997; Porter et al., 2005), the NOMINATE model (Poole & Rosenthal, 1985), and the IDEAL model (Jackman, 2001). For a review and comparison see Clinton et al. (2004) and Clinton & Jackman (2009).

One important challenge in the analysis of voting records from deliberative bodies is that they typically include varying degrees of missing observations. Usual analysis typically deal with missing values either by dropping votes that include them, or by inferring the missing values from the data as part of the inferential procedure. This approach involves assuming that the mechanism driving the appearance of missing observations is ignorable. However, the assumption of ignobility is typically made out of convenience and is often not justified in practice. Roughly speaking, a missing vote is ignorable if the mechanism through which a legislator decides whether to be absent is unrelated to the mechanism through which he decides the direction of his vote e.g., casting a"Aye" or "'Nay" vote (for a formal treatment, see for example Little & Rubin, 2002). In the case of legislative voting records, ignorability of missing values is not supported by empirical evidence, which suggests instead that non-responses in deliberative bodies are often the result of strategic decisions by the legislators (Cohen & Noll, 1991; Forgette, 1999; Noury, 2004; Kromer, 2005, 2008). Traditional explanations of mechanisms driving non-responses include indifference of marginal legislators to the outcome of the vote, the swing voter's curse, and the effect of competing principals, among others (Rosas & Shomer, 2008).

We are particularly interested in the role of competing principals in explaining non-responses by legislators. Competing principals refer to situations where the legislator's views collide with those of other political actors whose support is valued by the legislator, such as his constituency, his party or his coalition. In those situations it is possible that a legislator might decide to abstain rather than register a vote as a way to avoid marginalizing those political actors. This reasoning

[1] As a short hand, we refer to recorded votes on a particular proposition as a "bill", even though legislators often vote on motions, amendments, etc.

suggests a model in which the political ideology of the legislator and bill under consideration plays a key role in the probability of non-responses.

Following these ideas, we develop an extension of IDEAL that deals with non-ignorable missing values by jointly modeling the probability of non-response along with the probability of a positive vote (i.e., voting "Aye"). IDEAL is a Bayesian multivariate probit model (Chib & Greenberg, 1998); similar models have also been extensively used in the context of item response modeling (e.g., see Fox, 2010). IDEAL has been extended in a number of different directions, for example, Martin & Quinn (2002) incorporate a dynamic structure on the ideal points that helps understand the evolution of justices' opinions over time, while Hahn et al. (2012) describe a sparse factor model that can be used to explicitly test assumptions about the impact of different aspects of the policy space of the decisions of legislators. In a slightly different vein, Guo et al. (2011) have recently proposed the use of graphical models for discrete data to model the dependence in voting patterns across legislators.

The model we discuss in this paper is similar in spirit to those presented in Rosas & Shomer (2008) and Hans (2004) in that we jointly model the probability of non-response and the probability of a positive vote through the use of latent factors that affect both probabilities. Like Hans (2004), but unlike Rosas & Shomer (2008), we model the impact of ideology (as derived from the legislator's voting pattern) on the likelihood of non-response, rather than the other way around. However, we measure the impact of ideology on non-responses not only through the legislator's ideal points, but also through bill's ability to discriminate among legislators with extreme ideological positions. In one-dimensional spatial models of legislative behavior where ideological positions usually correlates strongly with party membership, this translates into models where not only the degree of partisanship of the legislators but also the level of partisanship of the bills can have an effect on the likelihood of non-response. As our examples illustrate, each of these factors can have varying importance in different political systems. Further distinguishing the present work, we explicitly incorporate model selection priors into our formulation, which allow us to formally test different hypotheses about the effect of policy positions on specific legislators and bills.

The rest of the paper is organized as follows. Section 2 reviews the basic formulation of Jackman (2001) and introduces our extension to modeling non-ignorable non-response. Section 3 describes our approach to computation based on Markov chain Monte Carlo algorithms. Section 4 presents two illustrations, one focusing on a series of votes about electoral reform in 19th century Britain, and one focusing on a full session of the 108th U.S. Congress. Finally, Section 5 discusses some limitation of our approach and future directions for research.

2 Model for revealed preferences in the presence of missing data

2.1 Modeling voting behavior

Let $y_{i,j}$ be the vote of legislator $i = 1, \dots, I$ on bill $j = 1, \dots, J$, where $y_{i,j} = 1$ if the vote is positive ("Aye") and $y_{i,j} = 0$ if the vote is negative ("Nay"). Assume first that all votes are observed for all legislators. We assume that, given latent parameters μ_j, α_j and β_i, the vote of legislator i in bill j follows a Bernoulli distribution with probability $\theta_{i,j}$, where

$$\theta_{i,j} = \mathsf{Pr}(y_{i,j} = 1 \mid \mu_j, \alpha_j, \beta_i) = \Phi\left(\mu_j + \alpha_j' \beta_i\right) \tag{1}$$

and Φ denotes the cumulative distribution function of the standard normal distribution. In this specification, the parameter μ_j represents the baseline probability of a positive vote for bill j, while the parameters α_j and β_i represent the position of bill j and legislator i in a latent "policy space" that is a subset of $\mathbb{R}^d$. In this model, it is the relative position of a given bill with respect to that of a given legislator's ideal point that controls the probability that such legislator will vote in favor of such bill. Indeed, the specification in (1) can be motivated from microeconomic arguments by assuming that each legislator must choose between two positions in policy space, ψ_j (corresponding to a positive vote on bill j) and ζ_j (corresponding to a negative vote), according to quadratic utility functions,

$$U_i(\psi_j) = -\left\|\psi_j - \beta_i\right\|^2 + \epsilon_{i,j}, \qquad\qquad U_i(\zeta_j) = -\left\|\zeta_j - \beta_i\right\|^2 + \nu_{i,j},$$

where $\epsilon_{i,j}$ and $\nu_{i,j}$ are stochastic shocks (associated, for example, to lobbying or party pressures) that might affect each specific vote in different ways.

Under this assumption, rational choice theory implies that legislator i will cast a positive vote on bill j (i.e., $y_{i,j} = 1$) if and only if $U_i(\psi_j) > U_i(\zeta_j)$. Assuming that the difference between the distribution of the difference between random utility shocks is normally distributed, $\epsilon_{i,j} - \nu_{i,j} \sim \mathsf{N}(0, \sigma_j^2)$, then

$$\mathsf{Pr}(y_{i,j} = 1 \mid \zeta_j, \psi_j, \sigma_j, \beta_i) = \Phi\left(\frac{\zeta_j'\zeta_j - \psi_j'\psi_j}{\sigma_j} - \frac{2(\zeta_j - \psi_j)'\beta_i}{\sigma_j}\right),$$

so that $\mu_j = (\zeta_j'\zeta_j - \psi_j'\psi_j)/\sigma_j$ and $\alpha_j = -2(\zeta_j - \psi_j)/\sigma_j$. Hence, if ψ_j and ζ_j are relatively close, then legislators tend to be indifferent between both policy positions no matter what their ideal points are. On the other hand, if ψ_j and ζ_j are well separated, then legislators have strong preferences and the vote will be clearly split.

It is important to note that, without additional constraints, the model just described is not identifiable because the utility functions are invariant to rotations, translations, rescalings and reflections of the policy space. A number of different approaches have been proposed in the literature to deal with the lack of identifiability, including restrictions on the mean and variances of the ideal points together with sign constrains in the discrimination parameters associated with the voting bills, as well as restrictions on the value of some legislators' ideal points. In this paper we follow this second approach and fix the ideal points associated with $d + 1$ legislators,[2] including those of the leaders of the main parties represented in the legislature (e.g., see Clinton et al., 2004). In a unidimensional model, this approach typically leads to ideal points that capture the position of the legislators in a liberal-conservative spectrum.

2.2 Modeling non-responses

In practice, records of legislative votes are seldom complete. These missing values might be due to explicit abstentions (in which the legislators explicitly affirm their choice not to vote), official absences (which correspond to situations in which the legislator is physically absent from the floor), or implicit abstentions (in which the legislator is present on the floor but fails to register a vote).

Although some legislative systems allow for explicit abstentions (for example, the New Zealand Parliament after 1995), differentiating between official absences and implicit abstentions might be difficult. Indeed, even is legislative systems where explicit abstentions are allowed, voting records might not differentiate between different types of abstentions (e.g., while explicit abstentions are technically possible in the US Senate at the discretion of the body by Section 2 of Rule XII, such abstentions are not recorded). Hence, in the sequel we do not differentiate among the different forms of non-responses, except in the cases in which information is available about the death or retirement of a given parliamentarian.

As mentioned in the introduction, a number of theories have been discussed in the literature to explain the strategic use of non-responses by legislators, and the structure of our model shall reflect the mechanism that is assumed to be driving the presence of missing values. For example, in the case of models that assume that missing values are driven by the indifference of legislators among

[2] Recall that d is the dimension of the policy space in which legislators' ideal points and bills reside.

outcomes, non-responses are more accurately modeled by extending the model in Section 2.1 from a binary to an ordinal response model in which a non-reponse represents a category that falls in between a positive and an negative vote. Since our focus in this paper is on modeling the effect of competing principals on legislators behavior, we believe that it is more appropriate to develop a joint model for the (hypothetical) voting record $y_{i,j}$ and the missing-data indicator $r_{i,j}$, where $r_{i,j} = 1$ if the vote of legislator i in bill j is missing, and $r_{i,j} = 0$ otherwise. When available, we also incorporate information about whether a given missing vote can be safely assumed to not be the result of a strategic decision driven by competing principals (as is often the case when votes are missing because of the death, retirement or permanent incapacitation of a legislator). This information is encoded through an additional set of indicators $s_{i,j}$ such that $s_{i,j} = 1$ when there is prior knowledge that the missing value is ignorable, and $s_{i,j} = 0$ otherwise. Our model assumes that such indicator is fixed in advance, with $s_{i,j} = 0$ for all i and j as default if no prior information is available.

In the case of missing values that are believed to be ignorable, a spatial model similar to the one discussed in Section 2.1 would be a natural choice, leading to

$$\mathsf{Pr}(r_{i,j} = 1 \mid s_{i,j} = 1, \eta_j, \lambda_j, \xi_i) = \Phi\left(\eta_j + \lambda_j' \xi_i\right), \tag{2}$$

where η_j represents the baseline propensity of legislators to miss votes in bill j, and λ_j and ξ_i represent, respectively, the positions of bill j and legislator i in a latent space that is a subset of $\mathbb{R}^q$ (note that this latent "missingness" space is different from the policy space discussed in Section 2.1). As before, the model can be justified using a utility-maximization argument. Furthermore, note that the missing observation is ignorable because the underlying utility function is independent of the actual vote $y_{i,j}$ and of the position of the legislator or the bill in policy space. Hence, the probability of non-response on a specific vote is independent of the probability of a positive vote. In this case, we ensure that the parameters are identifiable through our choice of priors for the latent features $\xi_1, \ldots \xi_I$ (see Section 2.3).

For non-responses that are assumed to be potentially non-ignorable (e.g., as the result of a strategic decision of a legislator subject to competing principals), the model described in (2) can be extended to incorporate the effect of both the position of the bill and the legislator's ideal point (both in policy space) on non-responses by setting

$$\mathsf{Pr}(r_{i,j} = 1 \mid s_{i,j} = 0, \eta_j, \lambda_j, \xi_i, \delta_j, \gamma_i, \alpha_j, \beta_i) = \Phi\left(\eta_j + \lambda_j' \xi_i + \alpha_j' \gamma_i + \delta_j' \beta_i\right). \tag{3}$$

In this specification the vectors γ_i and δ_j provide a link between our models for voting and non-response patterns. For example, in the case of the unidimensional policy space models where the position of the leader of the conservative party has been fixed to a positive value and the position of the leader of the liberal party has been fixed to a negative number, $\gamma_i > 0$ implies that legislator i is more likely to miss votes associated with bills supported by the conservative party, while $\delta_j > 0$ means that conservative legislators are more likely to miss votes on bill j. Similarly, note that setting $\gamma_i = 0$ for all $i = 1, \ldots, I$ and $\delta_j = 0$ for all $j = 1, \ldots, J$ brings us back to the case of ignorable missing values. Hence, by selecting appropriate priors for γ_i and δ_j we are able to explicitly test whether a given missing vote is ignorable or not under our model.

2.3 Prior distributions

The model is completed by specifying prior distributions for all model parameters. As is common practice in the context of Bayesian spatial models of legislative behavior, we focus on Gaussian and zero-inflated Gaussian priors, which lead a straightforward Markov chain Monte Carlo (MCMC) algorithm for posterior inference (see Section 3).

Consider first selecting priors for the parameters of the voting model described in Section 2.1. We assume that the baseline probabilities $\mu_1, \ldots, \mu_J$ are independent a priori with $\mu_j \sim \mathsf{N}(\rho_\mu, \omega_\mu^2)$ and $\rho_\mu \sim \mathsf{N}(0, 1)$ and $\omega_\mu^2 \sim \mathsf{IGam}(2, 1)$, where $\mathsf{IGam}(a, b)$ denotes the inverse Gamma distribution with a degrees of freedom and expectation $b/(a-1)$. Centering this prior distribution around zero reflects our belief that, a priori, legislators are equally likely to vote for or against any given bill. On the other hand, centering the prior variance around one is natural given that we work in a latent probit scale. Indeed, note that if $\alpha_j = 0$ (i.e., if the bill is not informative about the ideal points of the legislators), then setting $\mu_j \sim \mathsf{N}(0, 1)$ (which corresponds to substituting the expected values of the hyperparameters into the prior for μ_j) leads to $\theta_{i,j} \sim \mathsf{Uni}[0, 1]$ a priori. Similarly, the ideal points of the legislators are assigned a prior $\beta_i \sim \mathsf{N}(\rho_\beta, \Omega_\beta)$, with the d-dimensional mean vector ρ_β having a Gaussian prior with mean 0 and identity covariance matrix, and the $d \times d$ covariance matrix Ω_β being given an inverse-Wishart prior with $d + 1$ degrees of freedom and mean set to the identity matrix. Note that in the special case of a unidimensional policy space, this prior for Ω_β reduces to an inverse-Gamma prior with two degrees of freedom and unit mean. Finally, the entries of the vector α_j are given independent zero-inflated Gaussian distributions with

$$\alpha_{j,k} \sim \pi_{\alpha,k} \mathsf{N}(0, \omega_{\alpha,k}^2) + (1 - \pi_{\alpha,k}) \mathsf{I}_0$$

where I_0 denotes the degenerate measure at 0 and $\pi_{\alpha,k}$ corresponds to the prior probability that any given coefficient associated with the k-th dimension of the policy space is different to zero. A prior of this type allows us to explicitly contrast the hypotheses $\alpha_{j,k} = 0$ against $\alpha_{j,k} \neq 0$. Since $\alpha_{j,k} = 0$ corresponds to the assumption that bill j provides no information about the position of legislators on the k-th dimension of the policy space, this type of prior allows us to automatize the laborious and inexact process of manually excluding bills that are not informative about the policy space (e.g., because they correspond to unanimous votes). Furthermore, although we do not explore this direction in this paper, we note that $\alpha_{j,k} = 0$ for all $j = 1, \ldots, J$ corresponds to the hypothesis that the k-th dimension of the policy is not informative about legislators' behavior. Hence, this type of priors also allows us to potentially select the dimension of the policy space by selecting d moderately large and letting the model discard those dimensions of the policy space for which the data does not provide any information (for further discussion, see Hahn et al., 2012).

To account for the issues associated with simultaneously testing a very large number of hypothesis, we follow Scott & Berger (2003) and Scott & Berger (2010) and assign a prior distribution $\pi_{\alpha,k} \sim \mathsf{Beta}(\upsilon/d, 1)$. This prior is reminiscent of the stick breaking construction of the Indian Buffet process with parameter υ (Thibaux & Jordan, 2007; Teh et al., 2007), and it implies that the expected number of dimensions that are a priori important for a given legislator is υ. Since in this paper we do not attempt to estimate the dimension of the policy space we set $\upsilon = d$, which implies $\pi_{\alpha,k} \sim \mathsf{Uni}[0, 1]$ a priori.

Similar considerations come to mind when deriving prior distributions for the parameters of the non-response model described in Section 2.2. First, we assume that $\eta_1, \ldots, \eta_J$ are independent a priori with $\eta_j \sim \mathsf{N}(\rho_\eta, \omega_\eta^2)$, with $\rho_\eta \sim \mathsf{N}(0, 1)$ and $\omega_\eta^2 \sim \mathsf{IGam}(2, 1)$. Next, we assume that the ξ_is follow independent standard normal distribution for all legislators except the one with the most missing votes, which is assigned a standard normal distribution truncated to $(-\infty, 0)$. This truncation is introduced to ensure identifiability and, as suggested in Bafumi et al. (2005), we check whether it is supported by the data a posteriori. Finally, we set $\lambda_{j,k} \sim \pi_{\lambda,k}\mathsf{N}(0, \omega_{\lambda,k}^2) + (1 - \pi_{\lambda,k})\mathsf{I}_0$ with $\pi_{\lambda,k} \sim \mathsf{Uni}[0, 1]$.

To complete the model we need to specify priors on the vectors δ_j and γ_i that link the voting and missingness models. Since we are interested in assessing whether the position of bills and legislators in policy space indeed have an effect on the probability of non-responses, we also assign the entries of δ_j and γ_i zero inflated Guassian distributions,

$$\delta_{j,k} \sim \pi_{\delta,k}\mathsf{N}(0, \omega_{\delta,k}^2) + (1 - \pi_{\delta,k})\mathsf{I}_0, \qquad \gamma_{i,k} \sim \pi_{\gamma,k}\mathsf{N}(0, \omega_{\gamma,k}^2) + (1 - \pi_{\gamma,k})\mathsf{I}_0$$

with $\pi_{\delta,k} \sim \mathsf{Uni}[0, 1]$, $\pi_{\gamma,k} \sim \mathsf{Uni}[0, 1]$, $\omega_{\delta,k}^2 \sim \mathsf{IGam}(2, 1)$ and $\omega_{\gamma,k}^2 \sim \mathsf{IGam}(2, 1)$. As we mentioned before, setting $\delta_{j,k} = 0$ implies that the k-th dimension of the ideal point of the legislators has no impact on the probability of non-response for bill j (i.e., bill j is "immune" to partisanship effects along the k-th dimension of the policy space). On the other hand, $\gamma_{i,k} = 0$ implies that the level of partisanship of a given bill has no effect on whether legislator i will abstain on a given vote (i.e., whether legislator i regularly engages in strategic behavior). Furthermore, missing votes for which both $\gamma_{i,k} = 0$ and $\delta_{j,k} = 0$ for all k correspond to non-responses that are ignorable under our model.

3 Computation

The posterior distribution associated with our model, which is proportional to

$$\left\{ \prod_{k=1}^{d} p\left(\pi_{\alpha,k}\right) p\left(\pi_{\delta,k}\right) p\left(\pi_{\gamma,k}\right) p\left(\omega^2_{\alpha,k}\right), p\left(\omega^2_{\delta,k}\right), p\left(\omega^2_{\gamma,k}\right) \right\} \left\{ \prod_{k=1}^{q} p\left(\pi_{\lambda,k}\right) p\left(\omega^2_{\lambda,k}\right) \right\} p\left(\rho_\beta\right) p\left(\Omega_\beta\right)$$
$$p\left(\rho_\mu\right) p\left(\omega^2_\mu\right) p\left(\rho_\eta\right) p\left(\omega^2_\eta\right) \left\{ \prod_{j=1}^{J} p\left(\mu_j \mid \rho_\mu, \omega^2_\mu\right) p\left(\alpha_j \mid \{\pi_{\alpha,k}\}, \{\omega^2_{\alpha,k}\}\right) \right\} \left\{ \prod_{i=1}^{I} p\left(\beta_i \mid \rho_\beta, \Omega_\beta\right) \right\}$$
$$\left\{ \prod_{i=1}^{I} \prod_{j=1}^{J} p\left(y_{i,j} \mid \mu_{i,j}, \alpha_j, \beta_i\right) \right\} \left\{ \prod_{j=1}^{J} p\left(\eta_j \mid \rho_\eta, \omega^2_\eta\right) p\left(\lambda_j \mid \{\pi_{\lambda,k}\}, \{\omega^2_{\lambda,k}\}\right) p\left(\delta_j \mid \{\pi_{\delta,k}\}, \{\omega^2_{\delta,k}\}\right) \right\}$$
$$\left\{ \prod_{i=1}^{I} p\left(\xi_i\right) p\left(\gamma_i \mid \{\pi_{\gamma,k}\}, \{\omega^2_{\gamma,k}\}\right) \right\} \left\{ \prod_{i=1}^{I} \prod_{j=1}^{J} p\left(r_{i,j} \mid s_{i,j}, \eta_{i,j}, \lambda_j, \xi_i, \delta_j, \gamma_i, \alpha_j, \beta_i\right) \right\}, \quad (4)$$

is too complex to be treated analytically. Therefore, we resort to Markov chain Monte Carlo (MCMC) algorithms (Robert & Casella, 2005), which allow us to simulate a dependent sequence of random draws from the target posterior distribution. Given initial values for the parameters, the MCMC algorithm successively updates these values using the full conditional distributions derived from (4). Standard Markov chain theory ensures that, after an appropriate burn-in, the values of the parameters generated by the algorithm are approximately distributed according to the posterior distribution.

As commonly done for probit models, we employ the data-augmentation scheme described in Albert & Chib (1993). More specifically, we introduce auxiliary variables $z_{i,j} \sim \mathsf{N}(\mu_j + \alpha'_j \beta_i, 1)$ and $w_{i,j} \sim \mathsf{N}(\eta_j + \lambda'_j \xi_i + \alpha'_j \gamma_i + \delta'_j \beta_i, 1)$ such that $y_{i,j} = 1$ if and only if $z_{i,j} \geq 0$ and $r_{i,j} = 1$ if and only if $w_{i,j} \geq 0$. Conditionally on these auxiliary variables, the full conditional distributions for each μ_j, α_j, β_i, η_j, λ_j, ξ_i, γ_i, and δ_j are Gaussian, as are the full conditionals for ϖ and ρ. Similarly, the full conditional distributions for the variances $\{\omega_{\alpha,k}\}$, $\{\omega_{\lambda,k}\}$, $\{\omega_{\delta,k}\}$ and $\{\omega_{\gamma,k}\}$

follow inverse Gamma distributions, the full conditional distribution for the covariance matrix Ω_β follows an inverse-Wishart distribution, and the prior probabilities $\{\pi_{\alpha,k}\}$, $\{\pi_{\lambda,k}\}$, $\{\pi_{\delta,k}\}$ and $\{\pi_{\gamma,k}\}$ follow conditionally independent Beta full conditional posterior distributions. On the other hand, given all other parameters in the model, the latent variables $\{z_{i,j}\}$ and $\{w_{i,j}\}$ follow truncated Gaussian distributions, with the truncation region being determined by the values of $y_{i,j}$ and $r_{i,j}$, respectively. Furthermore, the identifiability constraints on the policy space (recall that we fixed the position of $d+1$ legislators) are implemented through a parameter expansion approach (Bafumi et al., 2005; Liu & Wu, 1998). More specifically, we first sample unconstrained ("expanded") parameters $\{\mu_j^*\}$, $\{\alpha_j^*\}$ and $\{\beta_i^*\}$, and the constraints are later enforced by applying an appropriate linear transformation. Details of the sampling algorithm for the case $d = 1$ and $q = 1$ (which correspond to the model used in our illustrations) are presented in Appendix A.

Given a sample from the previous Markov chain Monte Carlo algorithm,

$$\Big(\{z_{i,j}^{(b)}\}, \{w_{i,j}^{(b)}\}, \{\mu_j^{(b)}\}, \{\alpha_j^{(b)}\}, \{\beta_i^{(b)}\}, \{\eta_j^{(b)}\}, \{\lambda_j^{(b)}\}, \{\xi_i^{(b)}\}, \{\gamma_i^{(b)}\}, \{\delta_j^{(b)}\}, \rho_\mu^{(b)}, \omega_\mu^{(b)}, \{\omega_{\alpha,k}^{(b)}\},$$
$$\rho_\beta^{(b)}, \Omega_\beta^{(b)}, \rho_\eta^{(b)}, \omega_\eta^{(b)}, \{\omega_{\lambda,k}^{(b)}\}, \{\omega_{\delta,k}^{(b)}\}, \{\omega_{\gamma,k}^{(b)}\}, \{\pi_{\alpha,k}^{(b)}\}, \{\pi_{\lambda,k}^{(b)}\}, \{\pi_{\delta,k}^{(b)}\}, \{\pi_{\gamma,k}^{(b)}\}\Big), \quad b = 1, \ldots, B,$$

are obtained after an appropriate burn-in period, point and interval estimates for model parameters can be easily obtained by computing the empirical mean and/or the empirical quantiles of the posterior distribution. For example, in order to assess the relative position of legislators on each of the dimensions of the underlying policy space we can approximate the expected rank of each legislator as the empirical mean of the rank of $\beta_{i,k}^{(b)}$ across MCMC samples. Similarly, to assess the variability associated with those ranks we can compute the 2.5% and 97.5% quantiles of the ranks of $\beta_{i,k}^{(b)}$ across MCMC samples.

4 Illustrations

This section illustrates our approach to inference of legislators' revealed preferences in the presence of non-responses. For our illustrations we focus on unidimensional spatial models where $d = 1$ and $q = 1$. Hence, to identify our voting models we fix the values of the ideal points of the leaders of the two main parties in the corresponding legislative body. On the other hand, to identify the non-response model we constraint the sign of the latent factor associated with (one of) the individual(s) with the most abstentions to be negative.

All inferences presented below are based on 40,000 samples from our algorithm obtained after a burn-in period of 4,000 iterations. Convergence of the algorithm was assessed using the multi-chain approach discussed in Gelman & Rubin (1992); this method provided no evidence of lack of convergence. Sensitivity of the model to our prior specification is explored by changing the prior means for the variances of the latent factors and loadings. In both illustrations inferences appear to be robust to moderate changes in these parameters.

4.1 Electoral Reform in Britain, 1866 and 1867

Our first illustration focuses on legislative votes in the UK House of Commons concerning electoral reform, culminating in the Second Reform Act of 1867.[3] The Second Reform Act introduced a series of electoral reforms that extended the franchise to a large part of the urban male working class in England and Wales. Interestingly, electoral reforms were introduced by a majority-led Liberal government in 1866 which failed, causing a change in government. A minority-led Conservative government then introduced and successfully passed similar legislation in 1867.[4] Member's votes on reform-related legislation are recorded in divisions. Members' votes on reform-related divisions were obtained from *Parliamentary Papers*. A division is identified as reform-related if it took place during debate on reform. The data is taken from *Hansard*, which serves as an official record of members debates and speeches and has since the beginning of the 19th century. In total we analyze 10 divisions on the reform bill in 1866 and 49 divisions in 1867. Of these 59 divisions being considered, 39 failed (including one that failed by a single vote, and one that failed by 546) and only 20 passed. While political parties, and party membership, is not as well-defined as in modernity, MP's party is taken from parliamentary registers of the period (*Dod's Parliamentary Companion*; Mair, 1867).

Narrative accounts of the struggle for electoral reform in the Victorian House of Commons, and the passage of the Second Reform Act in 1867 include Smith (1966), Cox (1868) and Cowling (2005). Himmelfarb (1966) emphasizes partisan competition, above all, as a key explanatory factor in the eventual passage. However, the role of abstentions, which are substantial, has not been studied. Indeed, an interesting feature of this dataset is the high level of missing votes and high variability in turnout. About 10% of the 712 MPs (including members the House of Commons and of the House of Lords) missed at least 80% of the votes, and 20 participated in none of the votes (these 20 legislators were excluded from our analysis, leaving a sample size containing the 692

[3]We are grateful to Andrew Reeves for making available voting data from this period.

[4]For more on the factors explaining the failure in 1866 and success in 1867, see Moser & Reeves (January 2012).

MPs on the 59 divisions mentioned before).

We begin our analysis by presenting in Figure 1 the posterior mean and 95% credible intervals for the rank of legislators' ideal points under our competing-principals model. To estimate these ideal points we fixed the position of the Conservative leader, Benjamin Disraeli (fixed to $+1$), and that of the Liberal leader, William Gladstone (fixed to -1). As expected, the ideal points largely reflect party membership, with Liberals mostly on one end of the political spectrum, and Conservatives mostly on the other. However, the model reveals some surprises such as the position of Edward Cardwell, who appears to be the left-most member of the Conservative party, voting with the mainstream members of the Liberal party. Similarly, the positions of C. H. Mills and John Peel, who are technically members of the Conservative party, exhibit a voting records that position them well within the ranks of the Liberal party, while six members of the Liberal party (G. H. Heathcote, H. J. Selwin, B. Woodd, A. P. Clinton, H. W. Schneider, and A. H. A. Anson), have voting records that put them squarely within the ranks of the Conservative party, at least for the purposes of voting on reform. It is also worthwhile noting that there is a substantial level of uncertainty associated with the legislators' ideal points, particularly for those legislators with a large numbers of abstentions. Furthermore, the comparison presented in Figure 2 suggests that the estimates of the legislator's position in policy space can be substantially affected by the effects of competing principals, particularly when uncertainty of their position is large. For example, A. W. Young and H. W. Schneider, both Liberal MPs on paper, appear to be much more conservative when the pattern of abstentions is accounted for. Further, Figure 2 suggests estimates of Conservative's ideology is substially enhanced by incorporating missingness *vis-a-vis* our competing principles model (as seen by the cloud in the upper right of Figure 2).

Next, Figure 3 shows posterior means and posterior credible intervals for the division-specific parameters associated with the voting probabilities, along with posterior probabilities that specific α_js are equal to zero. We see divisions that are clearly favored by either mainstream liberals (located on the lower semiplane of Figure 3(a)) or mainstreams conservatives (located on the upper semiplane of of Figure 3(a)). On the other hand, the data provides evidence that two divisions (V124.67 and V107.67) contain limited information about legislators' ideal points, as the associated αs are likely zero. Division V124.67 (along with V123.67, for which $\mathsf{Pr}(\alpha_j = 0 \mid \text{Data}) \approx 0.38$) concerned representation of minorities and divided parties radically. Disreali and Gladstone, fiercely antagonistic towards each other as leaders of their respective parties, both opposed the motion while radicals Lowe and Mill - along with Earl Russell (the former Liberal Prime Minister) supported the motion. Referring to the eventual defeat of the motion (V124.67), a fellow MP and historian, Homersham Cox, commented "Of all the questions brought before Parliament during the discussion of the Reform Bill, there was no other debated so independently of the ordinary distinctions of parties." (Cox, 1868, pg 271)[5]. On the other hand, Division V107.67 concerned a one pound fine for neglect of overseers to collect and make available certain tax information. This Division attracted little attention from both MPs (it counts among the divisions with the fewest number of MPs voting – only 375) and historians.

Figure 4 presents the posterior mean and 95% credible intervals for the rank of the legislators based on $\xi_1, \ldots, \xi_{692}$, which capture the overall tendency of MPs to abstain from any given vote. To fully identify these parameters we constrained the latent factor associated with A. W. Young to be negative. This constraint seems to be satisfied a posteriori, as the 95% posterior credible interval of ξ_i for A. W. Young is $(-3.666, -1.909)$. Note that, although baseline abstentions are not driven by party affiliation in the same way that MPs ideal points are, there is a group of Liberal MPs concentrated at the top of the rankings. This suggests that an important faction of the Liberal party tended to be absent from most of the debate, suggesting a possible explanation for the failure to pass the reform in 1866 under a Liberal MP, and the success in 1867 under a more disciplined Conservative coalition.

Figure 5 presents posterior means and posterior credible intervals for the division-specific parameters $\{\eta_j\}$ and $\{\lambda_j\}$ associated with the pattern of missing votes, along with posterior probabilities that specific λ_js are equal to zero. The fact that the value of η_j is positive for some divisions agrees with our initial observation about the large number of missing observations. On the other hand, the fact that values of the λ_js are positive is consistent with imposing a negative constraint on the sign of the latent factor associated with the legislator with the most missing votes. Finally, note that in this setting all of the divisions seem to provide information about the likelihood that MPs randomly miss any given vote.

Finally, we proceed to evaluate the effect of MPs' ideology and location of divisions in policy space on the likelihood of missing votes. Figure 6 shows that there is little evidence in the data to support the claim that individual MPs tend to avoid particularly partisan divisions (the largest value for $\mathsf{Pr}(\gamma_i \neq 0 \mid \text{Data})$ corresponds to A. F. Grant, but it is only around 0.07). However, the results presented in Figure 7 suggest that extremely partisan parliamentarians do tend to miss a higher proportion of votes in a group of fifteen divisions (V48.66, V46.67, V158.67, V56.66, V62.66, V57.66, V109.67, V87.67, V157.67, V60.66, V154.67, V155.67, V39.67, V50.67 and

[5]The motion was eventually defeated, though an almost identical amendment was made by Lord Cairns and passed in the House of Lords.

V37.66). Except for V46.67, V87.67 and V50.67, the posterior mean of the coefficients δ_j are positive in all these divisions (see Figure 8). This implies that in most cases Conservative MPs (who, overall, have lower abstention rates than Liberal MPs) are more likely to miss votes on these divisions. These divisions are notable for at least two reasons. First, divisions taken late in 1867 were primarily on amendments made to the reform bill by the House of Lords (under Conservative leadership) most of which passed, but were unpopular in the House of Commons. It is hence noteworthy that ideologically conservative MPs in the House of Commons were prone to abstain from voting in such divisions. Next, it is interesting to note that among bills for which Conservatives are likely to abstain (i.e., bills for which $\delta_j > 0$), a disproportionate number occurred in 1866 and dealt with a technical issue of suffrage qualification known as the "rateable/rated" debate. While both technical and contentious, this aspect of reform dealt with the details of assessing financial requirements of potential voters. As some historians have noted, this aspect of the reform debate was poorly understood by most MPs and contentious among Liberals (Cox, 1868; McLean, 2001) (indeed, the vote effectively bringing down the Liberal government in 1866 was on just such an issue). It is telling, then, that these estimates support the interpretation of Conservatives essentially sitting out the ratable/rated debate, the implications of which few had any idea of in 1866 (when they were not in power) and rather letting the Liberals themselves "fight it out."

4.2 The second session of the 108th U.S. Senate

Our second illustration focuses on the analysis of the voting records from the second session of the 108th U.S. Senate (corresponding to the year 2004). During this session, the U.S. Senate was composed of 51 senators affiliated with the Republican party, 48 affiliated with the Democratic party and 1 independent who, for the most part, caucused with the Democrats (James Merrill Jeffords from Vermont). The voting record, which was obtained from the `THOMAS` database, `http://thomas.loc.gov/home/rollcallvotes.html`, includes 21,600 votes on 216 bills, 972 of which are missing (4.5% of the total). The distribution of missing votes varies widely; two senators (Kerry and Edwards, respectively the presidential and vice-presidential Democratic candidates in 2004) account for 323 of the missing votes (almost a third of the total), while 59 senators missed no more than four votes.

Figure 9 presents the posterior mean and 95% credible intervals for the rank of the Senators in policy space. To estimate these ideal points we fixed the positions of the Republican majority leader, Senator Frist from Tennessee (which was set to $+1$), and that of the Democratic minority leader, Senator Daschle from South Dakota (which was set to -1). As before, the ideal points of legislators in this unidimensional policy space largely reflect the effect of partisanship. The main exception is Senator Miller from Georgia who, despite technically being a Democrat, has a voting record that is closer to the most extreme Republican senators. In retrospect, this result is not surprising; Senator Miller backed Republican President George W. Bush over Democratic nominee John Kerry in the 2004 presidential election, and since 2003 frequently criticized the Democratic Party. Other interesting observations include the characterization of Senator Nelson from Nebraska as by far the most conservative mainstream Democrat in Congress, and the characterization of Senators Chafee from Rhode Island and Specter from Pennsylvania as well as Senators Snowe and Collins from Maine as the most centrist Republicans. Finally, we also highlight the high level of uncertainty associated with the rank of senator Kerry, which is most likely due to the large number of votes he missed during this period.

Figure 10 presents posterior means and posterior credible intervals for the bill-specific parameters associated with the voting probabilities, along with posterior probabilities that specific α_js are equal to zero. Bills that appear on the top half of Figure 10(a) (i.e., bills for which $\mathsf{Pr}(\alpha_j > 0 \mid \text{Data})$ is high) are Republican-leaning bills, while those on the bottom half are Democratic-leaning bills. Also, note that there is also a substantial number of bills that seem to provide limited (but non-negligible) information about the Senators. To better understand these results, we present in Figure 11 dotplots of the proportion of "yes" votes associated with 50 informative bills (bills with low values of $\mathsf{Pr}(\alpha_j = 0 \mid \text{Data})$) in the top panel and 50 uninformative bills (those with high values of $\mathsf{Pr}(\alpha_j = 0 \mid \text{Data})$) in the bottom panel. As would be expected, bills for which $\mathsf{Pr}(\alpha_j = 0 \mid \text{Data})$ is very small (i.e., bills that are highly informative about legislator's ideal points) are clearly partisan bills which, for the most part, were passed by a narrow majority comprised mostly of members of just one party. On the other hand, bills for which $\mathsf{Pr}(\alpha_j = 0 \mid \text{Data})$ is large (i.e., bills that contain limited information about ideal points) tend to be bills for which the "Ayes" are evenly split among both parties. Note, however, that in no case $\mathsf{Pr}(\alpha = 0 \mid \text{Data}) > 0.8$ (not even in the case of virtually unanimous bills, which would typically be excluded from the analysis by experts). This is probably because all of the unanimous bills were approved with a few senators abstaining. Hence, although our model heavily down-weights the effect of unanimous bills, it will still take into account some of the information provided by them to estimate legislators' ideal points.

Figure 12 presents the posterior mean and 95% credible intervals for the rank of the legislators in the latent space associated with the abstentions (i.e., "missingness" space). To estimate the positions of the Senators in this space we constrained the sign of the position of Senator Kerry (which had by far the largest number of abstentions) to be negative (again, the negative constraint seems to be innocuous, the 95% posterior credible interval for the ideal point of Senator Kerry is $(-1.21, -0.45)$). Not surprisingly, Democrat senators Kerry and Edwards (with 194 and 127 missing votes respectively) occupy the top two spots in this ranking, followed by Senators Campbell (41 missing votes) and Akaka (36), with very little uncertainty associated with their rankings. However, the absence of Senator Johnson (Democrat from South Dakota) from the top-five list (his mean rank is 13.11, with a 95% credible interval $(5, 27)$) is somewhat surprising, as he missed 37 votes (more than Akaka and almost as many as Campbell). This is because his abstentions seem to be driven by strategic decisions rather than just a high baseline rate of abstentions (see discussion below).

We proceed now to evaluate the effect of the position of Senators and bills in policy space on the likelihood of missing votes. Figure 13 shows the probability that the voting pattern of the different Senators is affected by the level of partisanship of the bill under consideration. Figure 13 provides very strong evidence that Democratic Senators Johnson (South Dakota) and Reid (Nevada) tend to preferentially miss votes on Democratic-leaning bills (the posterior means for the γ coefficients are -0.418 for Johnson and -0.333 for Reid, with corresponding 95% posterior credible intervals of $(-0.583, -0.277)$ and $(-0.495, -0.193)$), while providing no substantial evidence that any other Senator engages in strategic behavior (the next Senator in the list is Democrat Byrd from West Virginia, for which $\Pr(\gamma_j = 0 \mid \text{Data})$ is only 0.197). These results appear intuitively reasonable; both of Johnson and Reid represent conservative states whose constituency might not agree with all policy positions taken by the Democratic party. Hence, it makes sense that, for very controversial votes, these Senators might feel that not recording their vote will help them in future elections.

To better understand the estimates of the γ_js generated by our model, we consider the value of the point estimates of the α_js (which, as discussed before, measure the level of partisanship of the bills) associated with those bills in which each Senator abstained (see Figure 14). Note that five Senators at the bottom of the graph (three Democrats, Johnson, Reid and Byrd, and two Republicans, Capo and Voinovich) seem to have a preference for missing Democratic-leaning votes. However, only Johnson and Reid miss a number of votes that is large enough for the model to infer that their abstentions are not ignorable with any degree of statistical significance.

Finally, Figure 15 shows the probability that the voting pattern in different bills is affected by the ideology of Senators. Unlike the example of the Second Reform Act in the previous Section, there is no evidence that ideology affects any of the bills considered by the Senate during this period.

The results from this example highlights some advantages of our formulation over that of Hans (2004) and Rosas & Shomer (2008). In particular, note that if the discrimination parameters $\gamma_1, \ldots, \gamma_I$ were not included in the non-response model, there would be no evidence with which to investigate the claim that Senators in the U.S. Congress use abstentions strategically. As a final note, we also highlight the fact that, unlike our previous example, there is essentially no difference in the ranks of the Senators in policy space generated by this model and their ranks assuming that abstentions are ignorable.

5 Discussion

A key contribution of this paper is the ability to precisely identify votes for which abstentions are driven by legislator's competing principals. As our examples illustrate, this feature allows us to provide novel insights into the legislative process which can have impact both on the estimates of legislator's ideal points and our understanding of legislators' behavior.

By way of illustration, we showed two examples of different political settings in which missing values of recorded votes was accounted for. In the UK House of Commons, evidence of conservative MPs strategically abstaining is seen in the context of electoral reform (1866-67): partly in the relatively low-information setting of various aspects of electoral reform in 1866 and; in avoiding openly contradicting unpopular amendments supported in the Conservative-led House of Lords (1867). In the recent U.S. Senate, we saw no such role of strategic abstention. In both examples, our model allows for new inquiry including: testing explicitly the usually-assumed ignorability assumptions of missing votes; identifying legislators who strategically abstain – legislators whose ideology influences their presence in recorded votes.

One of the shortcoming of bilinear models such as the one described here is that it cannot capture interactions between legislator-specific and bill-specific features. For example, how liberal a legislator is might depend on whether the bill under considerations is a liberal or conservative bill. The same should be true for abstention patterns. This is particularly critical for understanding the behavior of "pivotal" legislators, which have a disproportionate impact on the final outcome.

A Details of the MCMC algorithm

We provide details of the MCMC algorithm described in Section 3 for the case where $d = 1$. Starting with initial values values for each of the parameters in model, these are updated by sampling from the following full conditional distributions:

1. To update the parameters associated with the policy space first generate samples for the expanded parameter $\{\mu_j^*\}$, $\{\alpha_j^*\}$ and $\{\beta_i^*\}$ using the following full conditionals:

   (a) For each $j = 1, \ldots, J$ sample $\mu_j^* \mid \cdots \sim \mathsf{N}(\hat{\rho}_{\mu,j}, \hat{\omega}^2_{\mu,j})$ where

$$\hat{\rho}_{\mu,j} = \frac{\frac{\rho_\mu}{\omega_\mu^2} + \sum_{i=1}^{I} \{z_{i,j} - \alpha_j \beta_i\}}{\frac{1}{\omega_\mu^2} + I} \qquad \text{and} \qquad \hat{\omega}^2_{\mu,j} = \frac{1}{\frac{1}{\omega_\mu^2} + I}.$$

   (b) For each $j = 1, \ldots, J$ sample α_j^* from a zero-inflated normal distribution $\alpha_j^* \mid \cdots \sim \hat{\pi}_{\alpha,j} \mathsf{N}(\hat{\rho}_{\alpha,j}, \hat{\omega}^2_{\alpha,j}) + \{1 - \hat{\pi}_{\alpha,j}\} \mathsf{I}_0$ where

$$\hat{\pi}_{\alpha,j} = \frac{\pi_\alpha \omega_\alpha^2}{\pi_\alpha \omega_\alpha^2 + (1 - \pi_\alpha) \omega^2_{\alpha,j} \exp\left\{ \frac{\rho^2_{\alpha,j}}{2\omega^2_{\alpha,j}} \right\}},$$

$$\hat{\rho}_{\alpha,j} = \frac{\sum_{i=1}^{I} \beta_i \{z_{i,j} - \mu_j\} + \sum_{i=1}^{I} \gamma_i \{w_{i,j} - \eta_j - \lambda_j \xi_i - \delta_j \beta_i\}}{\frac{1}{\omega_\alpha^2} + \sum_{i=1}^{I} \beta_i^2 + \sum_{i=1}^{I} \gamma_i^2},$$

   and

$$\hat{\omega}^2_{\alpha,j} = \frac{1}{\frac{1}{\omega_\alpha^2} + \sum_{i=1}^{I} \beta_i^2 + \sum_{i=1}^{I} \gamma_i^2}.$$

   (c) For each $i = 1, \ldots, I$ sample β_i^* from $\beta_i^* \mid \cdots \sim \mathsf{N}(\hat{\rho}_{\beta,i}, \hat{\omega}^2_{\beta,i})$ where

$$\hat{\rho}_{\beta,i} = \frac{\frac{\rho_\beta}{\omega_\beta^2} + \sum_{j=1}^{J} \alpha_j \{z_{i,j} - \mu_j\} + \sum_{j=1}^{J} \delta_j \{w_{i,j} - \eta_j - \lambda_j \xi_i - \gamma_i \alpha_j\}}{\frac{1}{\omega_\beta^2} + \sum_{j=1}^{J} \alpha_j^2 + \sum_{j=1}^{J} \delta_j^2},$$

   and

$$\hat{\omega}^2_{\beta,i} = \frac{1}{\frac{1}{\omega_\beta^2} + \sum_{j=1}^{J} \alpha_j^2 + \sum_{j=1}^{J} \delta_j^2}.$$

   Once these samples have been obtained, enforce the identifiability constraints by setting

$$\mu_j = \mu_j^* + \alpha_j^* \left(\frac{\beta^*_{s_2} - \beta^*_{s_1}}{2} \right), \quad \alpha_j = \alpha_j^* \left(\frac{\beta^*_{s_2} - \beta^*_{s_1}}{2} \right) \quad \text{and} \quad \beta_i = \frac{2\beta_i - \beta^*_{s_2} - \beta^*_{s_1}}{\beta^*_{s_2} - \beta^*_{s_1}}.$$

   where s_1 and s_2 are the indexes of the legislators whose ideal points are to be set to -1 and 1, respectively.

2. For each $i = 1, \dots, I$, sample γ_i from a zero-inflated normal distribution $\gamma_i \cdots \sim \hat{\pi}_{\gamma,i} \mathsf{N}(\hat{\rho}_{\gamma,i}, \hat{\omega}^2_{\gamma,i}) + \{1 - \hat{\pi}_{\gamma,i}\} \mathsf{I}_0$, where

$$\hat{\pi}_{\gamma,i} = \frac{\pi_\gamma \omega_\gamma^2}{\pi_\gamma \omega_\gamma^2 + (1-\pi_\gamma)\omega^2_{\gamma,j} \exp\left\{\frac{\rho^2_{\gamma,j}}{2\omega^2_{\gamma,j}}\right\}}, \qquad \hat{\rho}_{\gamma,i} = \frac{\sum_{j=1}^J \alpha_j \{w_{i,j} - \eta_j - \lambda_j \xi_i - \beta_i \delta_j\}}{\frac{1}{\omega_\gamma^2} + \sum_{j=1}^J \alpha_j^2},$$

and

$$\hat{\omega}^2_{\gamma,i} = \frac{1}{\frac{1}{\omega_\gamma^2} + \sum_{j=1}^J \alpha_j^2}.$$

3. For each $j = 1, \dots, J$, sample δ_j from a zero-inflated normal distribution $\delta_j \cdots \sim \hat{\pi}_{\delta,j} \mathsf{N}(\hat{\rho}_{\delta,j}, \hat{\omega}^2_{\delta,j}) + \{1 - \hat{\pi}_{\delta,j}\} \mathsf{I}_0$, where

$$\hat{\pi}_{\delta,j} = \frac{\pi_\delta \omega_\delta^2}{\pi_\delta \omega_\delta^2 + (1-\pi_\delta)\omega^2_{\delta,j} \exp\left\{\frac{\rho^2_{\delta,j}}{2\omega^2_{\delta,j}}\right\}}, \qquad \hat{\rho}_{\delta,j} = \frac{\sum_{i=1}^I \beta_i \{w_{i,j} - \eta_j - \lambda_j \xi_i - \alpha_j \gamma_i\}}{\frac{1}{\omega_\delta^2} + \sum_{i=1}^I \beta_i^2},$$

and

$$\hat{\omega}^2_{\delta,j} = \frac{1}{\frac{1}{\omega_\delta^2} + \sum_{i=1}^I \beta_i^2}.$$

4. For each $j = 1, \dots, J$ sample $\eta_j \mid \cdots \sim \mathsf{N}(\hat{\rho}_{\eta,j}, \hat{\omega}^2_{\eta,j})$ where

$$\hat{\rho}_{\eta,j} = \frac{\frac{\rho_\eta}{\omega_\eta^2} + \sum_{i=1}^I \{w_{i,j} - \lambda_j \xi_i - \alpha_j \gamma_i - \beta_i \delta_j\}}{\frac{1}{\omega_\eta^2} + I} \qquad \text{and} \qquad \hat{\omega}^2_{\eta,j} = \frac{1}{\frac{1}{\omega_\eta^2} + I}.$$

5. For each $j = 1, \dots, J$ sample λ_j from a zero-inflated normal distribution $\lambda_j \mid \cdots \sim \hat{\pi}_{\lambda,j} \mathsf{N}(\hat{\rho}_{\lambda,j}, \hat{\omega}^2_{\lambda,j}) + \{1 - \hat{\pi}_{\lambda,j}\} \mathsf{I}_0$ where

$$\hat{\pi}_{\lambda,j} = \frac{\pi_\lambda \omega_\lambda^2}{\pi_\lambda \omega_\lambda^2 + (1-\pi_\lambda)\omega^2_{\lambda,j} \exp\left\{\frac{\rho^2_{\lambda,j}}{2\omega^2_{\lambda,j}}\right\}}, \quad \hat{\rho}_{\lambda,j} = \frac{\sum_{i=1}^I \xi_i \{w_{i,j} - \eta_j - \alpha_j \gamma_i - \delta_j \beta_i\}}{\frac{1}{\omega_\lambda^2} + \sum_{i=1}^I \xi_i^2},$$

and

$$\hat{\omega}^2_{\lambda,j} = \frac{1}{\frac{1}{\omega_\lambda^2} + \sum_{i=1}^I \xi_i^2}.$$

6. For each $i = 1, \dots, I$ but $i \neq s_3$ (where s_3 denotes the index of the legislator with the most missing votes in the record), sample ξ_i from $\xi_i \mid \cdots \sim \mathsf{N}(\hat{\rho}_{\xi,i}, \hat{\omega}^2_{\xi,i})$ where

$$\hat{\rho}_{\xi,i} = \frac{\sum_{j=1}^J \lambda_j \{w_{i,j} - \eta_j - \delta_j \beta_i - \gamma_i \alpha_j\}}{1 + \sum_{j=1}^J \lambda_j^2},$$

and

$$\hat{\omega}^2_{\xi,i} = \frac{1}{1 + \sum_{j=1}^{J} \lambda_j^2}.$$

On the other hand, sample $\xi_{s_3} \mid \cdots \sim \mathsf{N}(\hat{\rho}_{\xi,s_3}, \hat{\omega}^2_{\xi,s_3})\mathsf{1}_{(0,1)}$

7. for $i = 1, \ldots, I$ and $j = 1, \ldots, J$, update the auxiliary variable $z_{i,j}$ by sampling from

$$z_{i,j} \mid \cdots \sim \begin{cases} \mathsf{N}(\mu_j + \alpha_j \beta_i, 1)\mathsf{I}_{(z_{i,j} \geq 0)} & y_{i,j} = 1, r_{i,j} = 0 \\ \mathsf{N}(\mu_j + \alpha_j \beta_i, 1)\mathsf{I}_{(z_{i,j} < 0)} & y_{i,j} = 0, r_{i,j} = 0 \\ \mathsf{N}(\mu_j + \alpha_j \beta_i, 1) & r_{i,j} = 1 \end{cases}$$

where $\mathsf{N}(a, b^2)\mathsf{I}_{\Omega}$ denotes the Gaussian distribution with mean a and variance b constrained to the set Ω.

8. for $i = 1, \ldots, I$ and $j = 1, \ldots, J$, update the auxiliary variable $w_{i,j}$ by sampling from

$$w_{i,j} \mid \cdots \sim \begin{cases} \mathsf{N}(\eta_j + \lambda_j \xi_i + \alpha_j \gamma_i + \delta_j \beta_i, 1)\mathsf{I}_{(z_{i,j} \geq 0)} & r_{i,j} = 1 \\ \mathsf{N}(\eta_j + \lambda_j \xi_i + \alpha_j \gamma_i + \delta_j \beta_i, 1)\mathsf{I}_{(z_{i,j} < 0)} & r_{i,j} = 0 \end{cases}$$

9. Update ρ_μ by sampling from

$$\rho_\mu \mid \cdots \sim \mathsf{N}\left(\frac{\frac{\sum_{j=1}^{J} \mu_j}{\omega_\mu^2}}{1 + \frac{J}{\omega_\mu^2}}, \frac{1}{1 + \frac{J}{\omega_\mu^2}} \right).$$

10. Update ω_μ^2 by sampling from

$$\omega_\mu^2 \mid \cdots \sim \mathsf{IGam}\left(2 + \frac{J}{2}, 1 + \frac{\sum_{j=1}^{J} (\mu_j - \rho_\mu)^2}{2} \right).$$

where $\mathsf{IGam}(a, b)$ denotes the inverse Gamma distribution with a degrees of freedom and mean $b/(a-1)$.

11. Update ρ_β by sampling from

$$\rho_\beta \mid \cdots \sim \mathsf{N}\left(\frac{\frac{\sum_{i=1}^{I} \beta_j}{\omega_\beta^2}}{1 + \frac{I}{\omega_\beta^2}}, \frac{1}{1 + \frac{I}{\omega_\beta^2}} \right).$$

12. Update ω_β^2 by sampling from

$$\omega_\beta^2 \mid \cdots \sim \mathsf{IGam}\left(2 + \frac{I}{2}, 1 + \frac{\sum_{i=1}^{I}(\beta_i - \rho_\beta)^2}{2}\right).$$

13. Update ω_α^2 by sampling from

$$\omega_\alpha^2 \mid \cdots \sim \mathsf{IGam}\left(2 + \frac{n_\alpha}{2}, 1 + \frac{\sum_{\{j:\alpha_j \neq 0\}} \alpha_j^2}{2}\right),$$

where $n_\alpha = \sum_{\{j:\alpha_j \neq 0\}} 1$ is the number of α_js that are different from zero.

14. Update ρ_η by sampling from

$$\rho_\eta \mid \cdots \sim \mathsf{N}\left(\frac{\frac{\sum_{j=1}^{J} \eta_j}{\omega_\eta^2}}{1 + \frac{J}{\omega_\eta^2}}, \frac{1}{1 + \frac{J}{\omega_\eta^2}}\right).$$

15. Update ω_η^2 by sampling from

$$\omega_\eta^2 \mid \cdots \sim \mathsf{IGam}\left(2 + \frac{J}{2}, 1 + \frac{\sum_{j=1}^{J}(\eta_j - \rho_\eta)^2}{2}\right).$$

16. Update ω_λ^2 by sampling from

$$\omega_\lambda^2 \mid \cdots \sim \mathsf{IGam}\left(2 + \frac{n_\lambda}{2}, 1 + \frac{\sum_{\{j:\lambda_j \neq 0\}} \lambda_j^2}{2}\right).$$

where $n_\lambda = \sum_{\{j:\lambda_j \neq 0\}} 1$ is the number of λ_js that are different from zero.

17. Update ω_γ^2 by sampling from

$$\omega_\gamma^2 \mid \cdots \sim \mathsf{IGam}\left(2 + \frac{n_\gamma}{2}, 1 + \frac{\sum_{\{i:\gamma_i \neq 0\}} \gamma_i^2}{2}\right).$$

where $n_\gamma = \sum_{\{i:\gamma_i \neq 0\}} 1$ is the number of γ_is that are different from zero.

18. Update ω_δ^2 by sampling from

$$\omega_\delta^2 \mid \cdots \sim \mathsf{IGam}\left(2 + \frac{n_\delta}{2}, 1 + \frac{\sum_{\{j:\delta_j \neq 0\}} \delta_j^2}{2}\right).$$

where $n_\delta = \sum_{\{j:\delta_j \neq 0\}} 1$ is the number of δ_js that are different from zero.

19. Update π_α by sampling from $\pi_\alpha \mid \cdots \sim \mathsf{Beta}(1 + n_\alpha, 1 + J - n_\alpha)$.

20. Update π_λ by sampling from $\pi_\lambda \mid \cdots \sim \mathsf{Beta}(1 + n_\lambda, 1 + J - n_\lambda)$.

21. Update π_γ by sampling from $\pi_\gamma \mid \cdots \sim \mathsf{Beta}(1 + n_\gamma, 1 + I - n_\gamma)$.

22. Update π_δ by sampling from $\pi_\delta \mid \cdots \sim \mathsf{Beta}(1 + n_\delta, 1 + J - n_\delta)$.

References

ALBERT, J. & CHIB, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association* **88**, 669–679.

BAFUMI, J., GELMAN, A., PARK, D. K. & KAPLAN, N. (2005). Practical issues in implementing and understanding Bayesian ideal point estimation. *Political Analysis* **13**, 171–187.

CHIB, S. & GREENBERG, E. (1998). Analysis of multivariate probit models. *Biometrika* **85**, 347–361.

CLINTON, J. & JACKMAN, S. (2009). To simulate or NOMINATE? *Legislative Studies Quarterly* **34**, 593–621.

CLINTON, J., JACKMAN, S. & RIVERS, D. (2004). The statistical analysis of roll call data. *American Political Sicence Review* **98**, 355–370.

COHEN, L. & NOLL, R. (1991). How to vote, thether to vote: Strategies for voting and bstaining in congressional roll calls. *Political Behavior* **13**, 97–127.

COWLING, M. (2005). *1867: Disraeli, Gladstone and revolution: the passing of the second Reform Bill*. Cambridge University Press.

COX, H. (1868). *A history of the reform bills of 1866 and 1867*. Longmans, Green, and co.

DAVIS, O., HINICH, M. J. & ORDESHOOK, P. C. (1970). An expository development of a mathematical model of the electoral process. *American Political Sicence Review* **64**, 426–248.

Dod's Parliamentary Companion (1865). London: Whittaker & Co.

ENELOW, J. M. & HINICH, M. J. (1984). *The Spatial Theory of Voting: An Introduction*. New York: Cambridge University Press.

FORGETTE, B., R. AN DSALA (1999). Conditional party government and member turnout on senate recorded votes, 1873-1935. *Journal of Politics* **61**, 467–484.

FOX, J.-P. (2010). *Bayesian Item Response Modeling: Theory and Applications*. Springer.

GELMAN, A. & RUBIN, D. (1992). Inferences from iterative simulation using multiple sequences. *Statistical Science* **7**, 457–472.

GUO, J., LEVINA, E., MICHAILIDIS, G. & ZHU, J. (2011). Estimating heterogeneous graphical models for discrete data with an application to roll call voting. Technical report, Department of Statistics, University of Michigan.

HAHN, P. R., CARVALHO, C. M. & SCOTT, J. G. (2012). A sparse factor analytic probit model for congressional voting patterns. *Applied Statistics* **61**.

HANS, N. (2004). The dog that didn't bark: Using abstention in estimating ideal points from roll call data. Poster presented at the annual conference of the Political Methodology Society, Palo Alto, CA.

HECKERMAN, J. J. & SNYDER, J. M. (1997). Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. *RAND Journal of Economics* **28**, S142–S189.

HIMMELFARB, G. (1966). The politics of democracy: The english reform act of 1867. *Journal of British Studies* **6**, 97–138.

JACKMAN, S. (2001). Multidimensional analysis of roll call data via Bayesian simulation: Identification, estimation, inferene and model checking. *Political Analysis* **9**, 227–241.

JENKKINS, S. (2006). The impact of party and ideology on roll-call voting in state legislatures. *Legislative Studies Quarterly* **31**, 235–257.

KROMER, M. K. (2005). *Determinants of abstention in the United States House of Representatives: An analysis of the 102nd through the 107th session*. Master's thesis, Louisiana State University, Baton Rouge, LA.

KROMER, M. K. (2008). *Legislative Participation in the American States: Determinants and Consequence*. Ph.D. thesis, Louisiana State University, Baton Rouge, LA.

LITTLE, R. J. A. & RUBIN, D. B. (2002). *Statistical analysis with missing data*. New York: Wiley, 2nd edition.

LIU, D. B., C. ND RUBIN & WU, Y. (1998). Parameter expansion to accelerate EM: The PX-EM algorithm. *Biometrika* **85**, 755–770.

MAIR, R. H. (1867). *Debrett's illustrated House of Commons and the Judicial Bench*. London: Dean and Son.

MARTIN, A. D. & QUINN, K. M. (2002). Dynamic ideal point estimation via markov chain monte carlo for the u.s. supreme court, 1953-1999. *Political Analysis* **10**, 134–153.

MCCARTY, M., POOLE, K. T. & ROSENTHAL, H. (2006). *Polarized America: the Dance of Ideology and Unequal Riches*. Cambridge: MIT Press.

MCLEAN, I. (2001). *Rational Choice and British Politics*. Oxford University Press.

MOSER, S. & REEVES, A. (January 2012). Taking the leap: Voting, rhetoric, and the determinants of electoral reform. Available at SSRN: http://dx.doi.org/10.2139/ssrn.1980338.

NOURY, A. (2004). Abstention in daylight: Strategic calculus of voting in the European parliament. *Public Choice* **212**, 179–211.

POOLE, K. T. & ROSENTHAL, H. (1985). A spatial model for legislative roll call analysis. *American Journal of Political Science* **29**, 357–384.

PORTER, M. A., MUCHA, P. J., NEWMAN, M. E. J. & WARMBRAND, C. M. (2005). A network analysis of committees in the u.s. house of representatives. *Proceedings of the National Academy of Sciences* **102**, 7057–7062.

ROBERT, C. P. & CASELLA, G. (2005). *Monte Carlo Statistical Methods*. Springer, second edition edition.

ROSAS, G. & SHOMER, Y. (2008). Models of nonresponse in legislative politics. *Legislative Studies Quarterly* **33**, 573–601.

SCHICKLER, E. (2000). Institutional change in the House of Representatives, 1867-1998: A test of partisan and ideological power balance models. *American Political Sicence Review* **44**, 187–205.

SCOTT, J. & BERGER, J. O. (2003). An exploration of aspects of Bayesian multiple testing. *Journal of Statistical Planning and Inference* **137**, 2144–2162.

SCOTT, J. G. & BERGER, J. O. (2010). Bayes and empirical-bayes multiplicity adjustment in the variable-selection problem bayes and empirical-bayes multiplicity adjustment in the variable-selection problem bayes and empirical-bayes multiplicity adjustment in the variable-selection problem. *Annals of Statistics* **38**, 2587–2619.

SMITH, F. B. (1966). *The making of the Second Reform Bill*. Cambridge U.P.

TEH, Y. W., GÖRÜR, D. & GHAHRAMANI, Z. (2007). Stick-breaking construction for the Indian buffet process. In *Proceedings of the Workshop on Artificial Intelligence and Statistics*.

THIBAUX, R. & JORDAN, M. (2007). Hierarchical beta processes and the indian buffet process. In *Proceedings of the 11th Conference on Artificial Intelligence and Statistics (AISTAT)*. Puerto Rico.

WRIGHT, G. C. & SCHAFFNER, B. F. (2002). The influence of party: Evidence from the state legislatures. *American Political Sicence Review* **96**, 367–379.

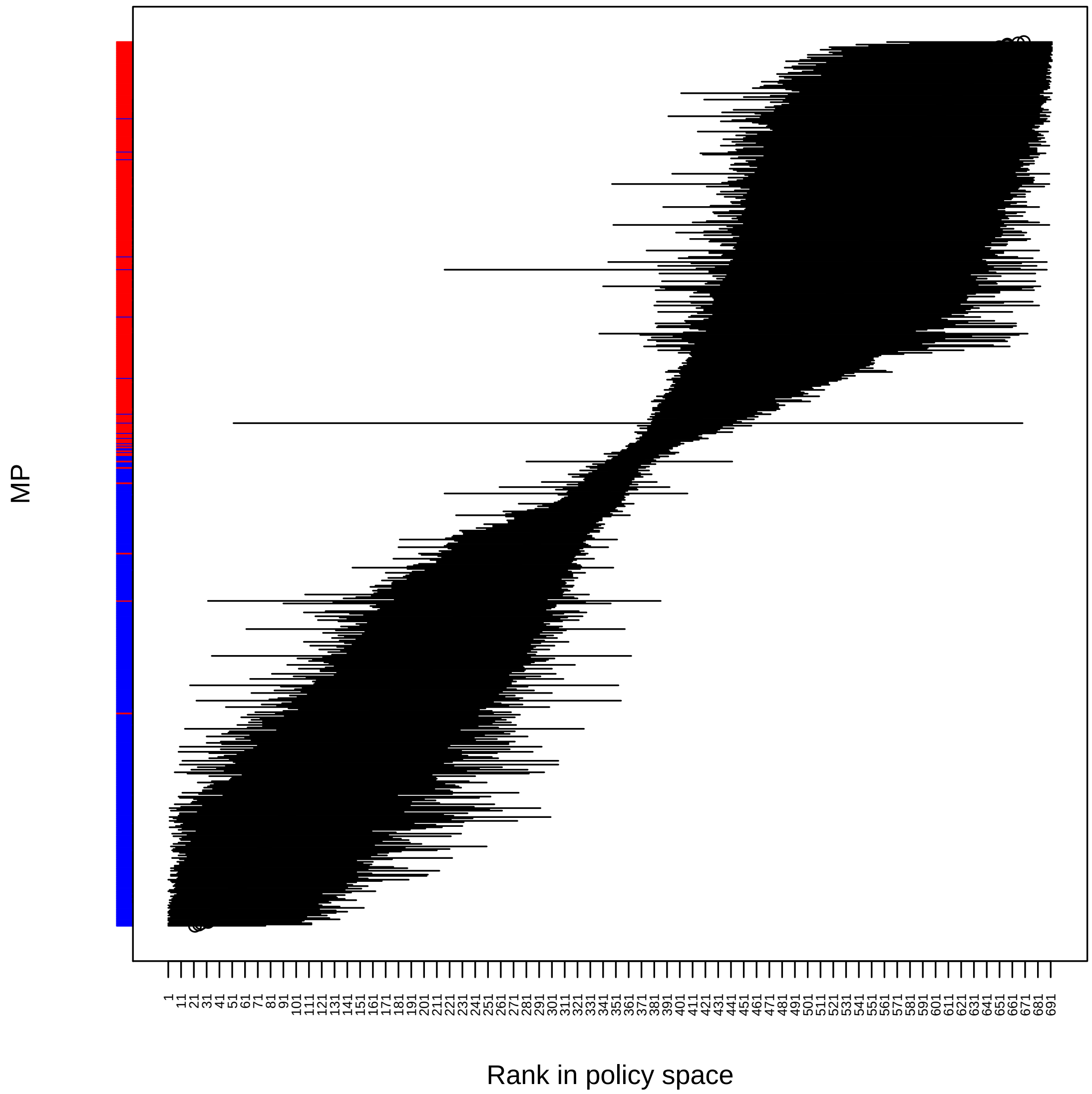


Figure 1: Posterior means and posterior credible intervals of British MP's ideal points based on voting during 1866-1867 reform debates in the House of Commons. Members of the Liberal party are identified in blue, while members of the Conservative party are depicted in red.

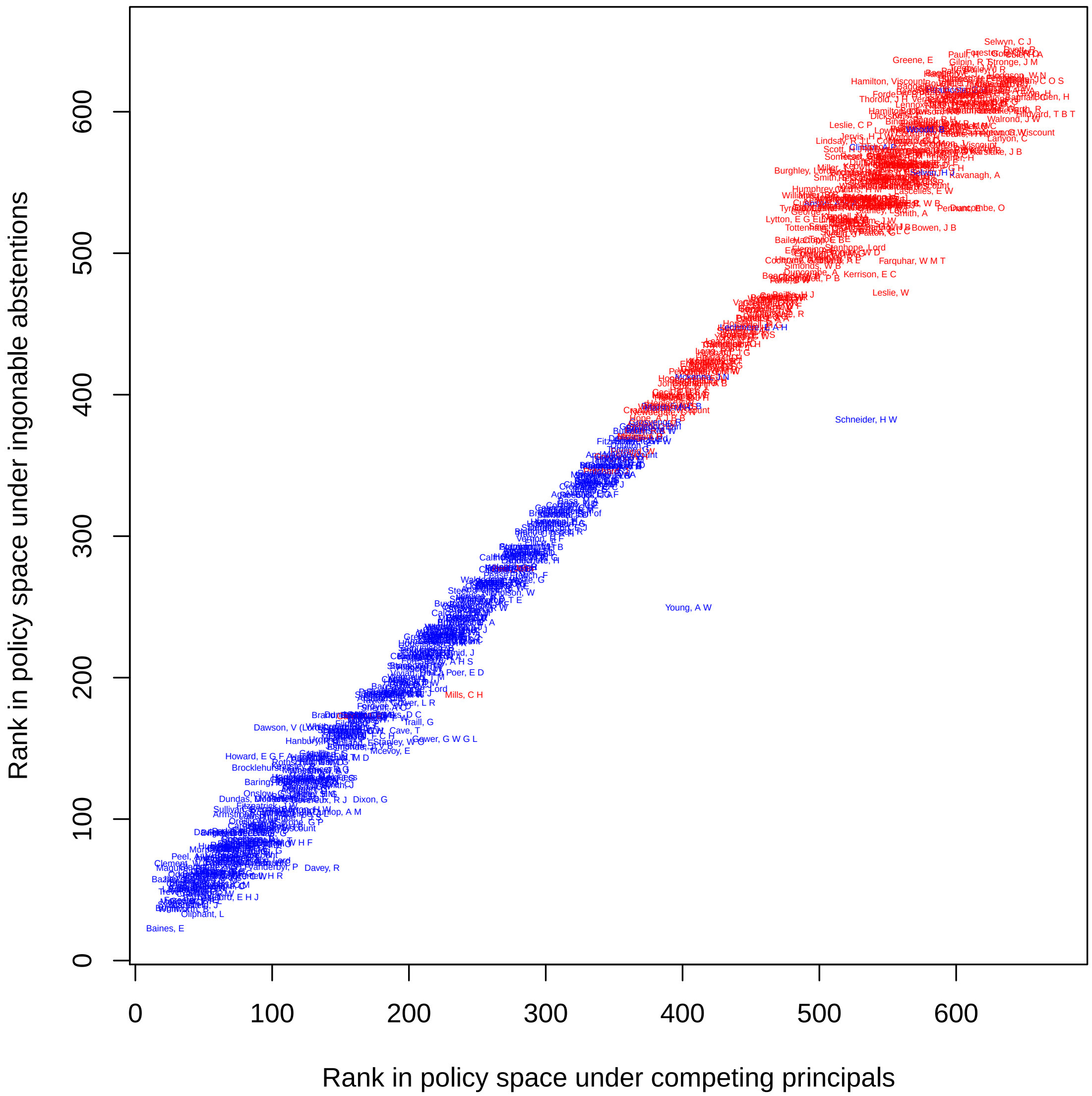
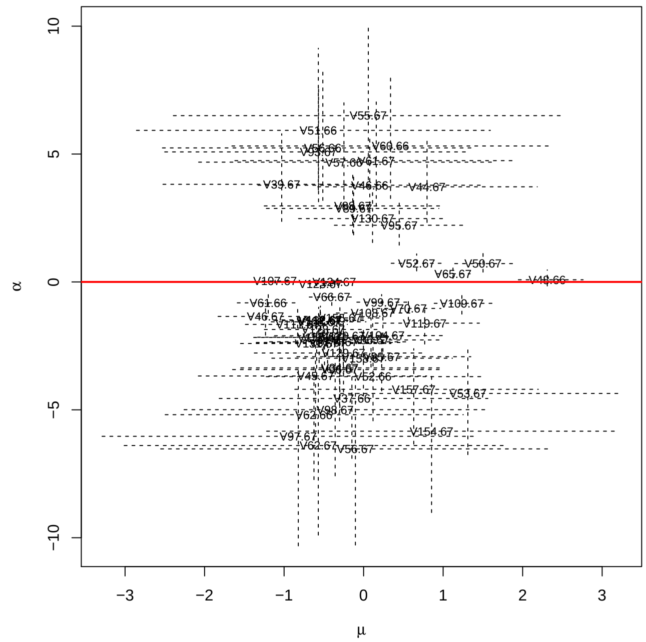
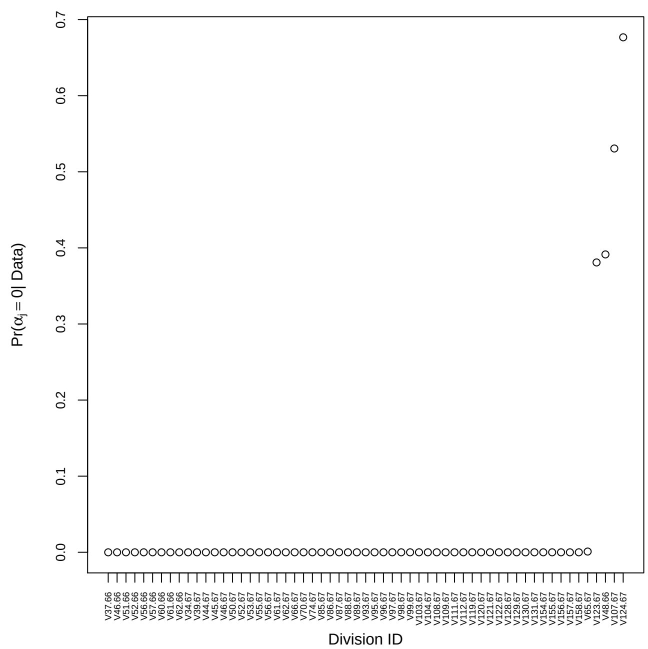


Figure 2: Posterior mean of ideal points generated by our model for competing principals against those generated by a model that assumes that abstentions are ignorable. Colors represent party membership, with red corresponding to members of the Conservative party, and blue representing Liberal MPs.



(a)



(b)

Figure 3: Panel (a) shows posterior means and posterior credible intervals for division-specific parameters associated with the voting patterns of the British House of Commons electoral reform deliberation of 1866-1867. Panel (b) shows the posterior probability that a given division is not informative about the ideal points of members.

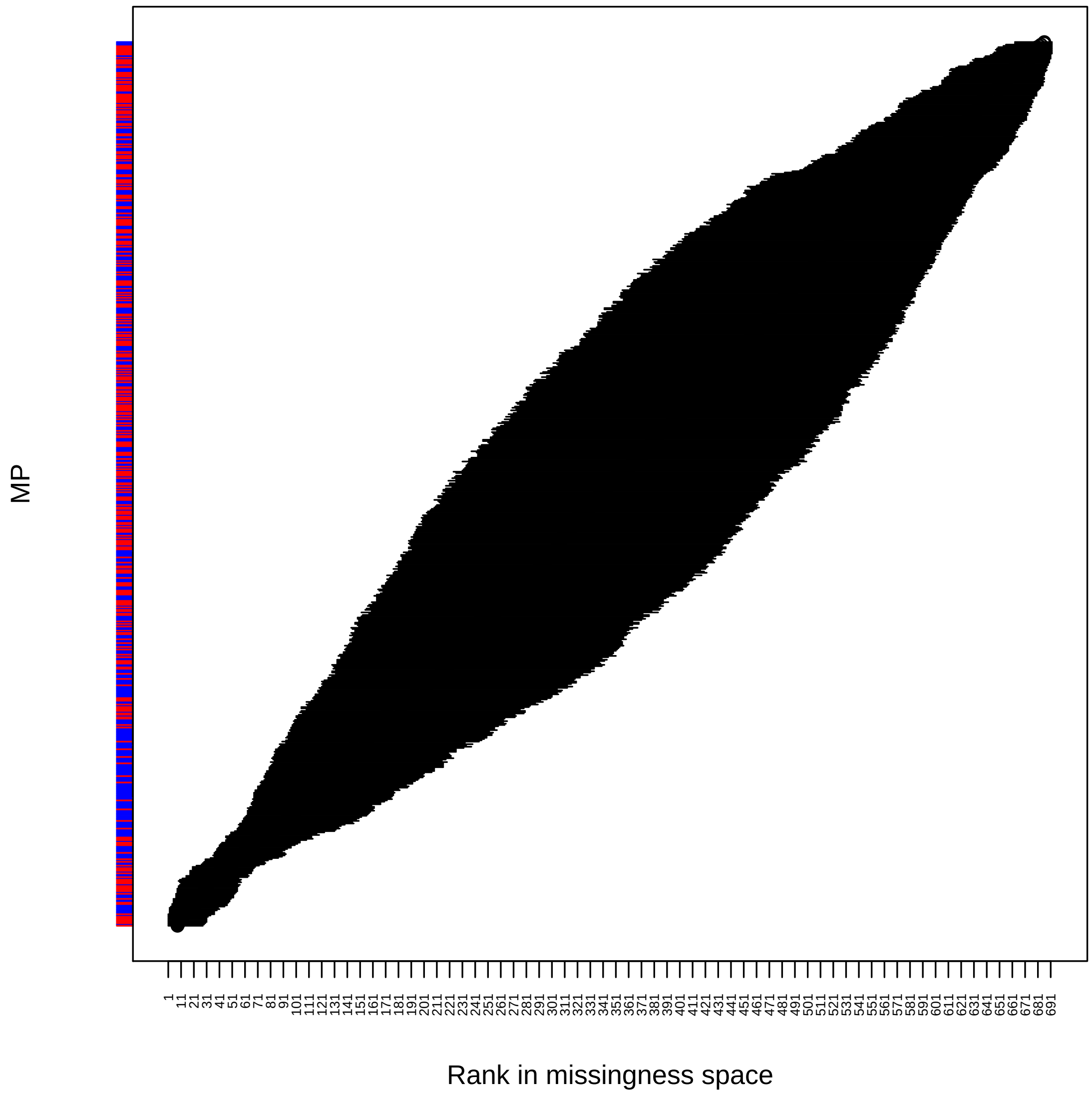
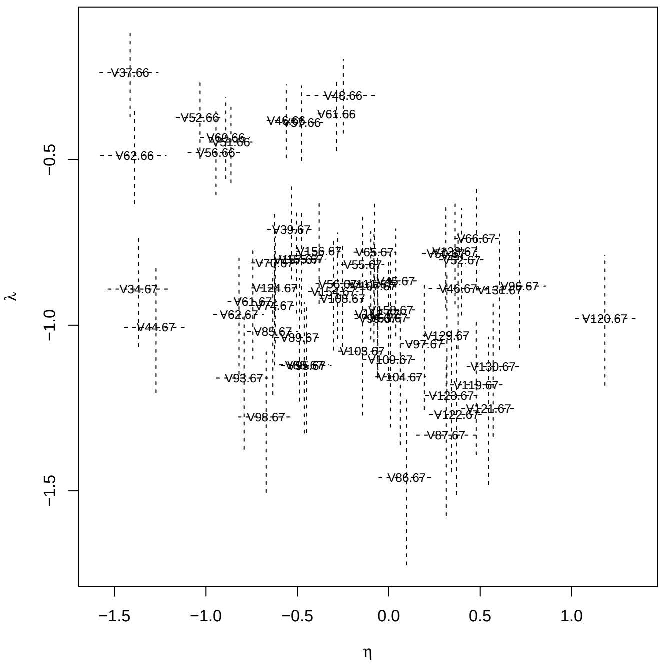
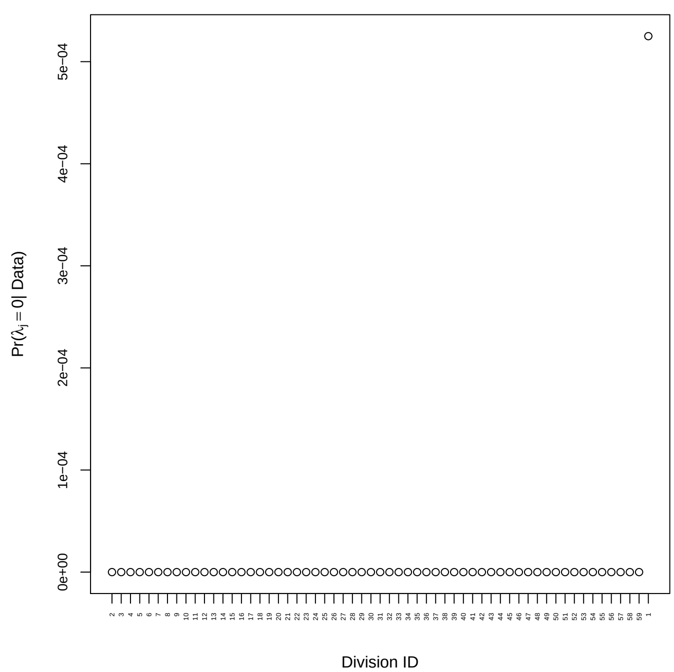


Figure 4: Posterior means and posterior credible intervals for the positions in missingness space for British MPs deliberating electoral reform, 1866-1867. Members of the Liberal party are identified in blue, while members of the Conservative party are identified in red.



(a)



(b)

Figure 5: Panel (a) shows posterior means and posterior credible intervals for division-specific parameters associated with the pattern of missing votes in the British voting reform. Panel (b) shows the posterior probability that a given division is not informative about the probability that a MP randomly misses a vote.

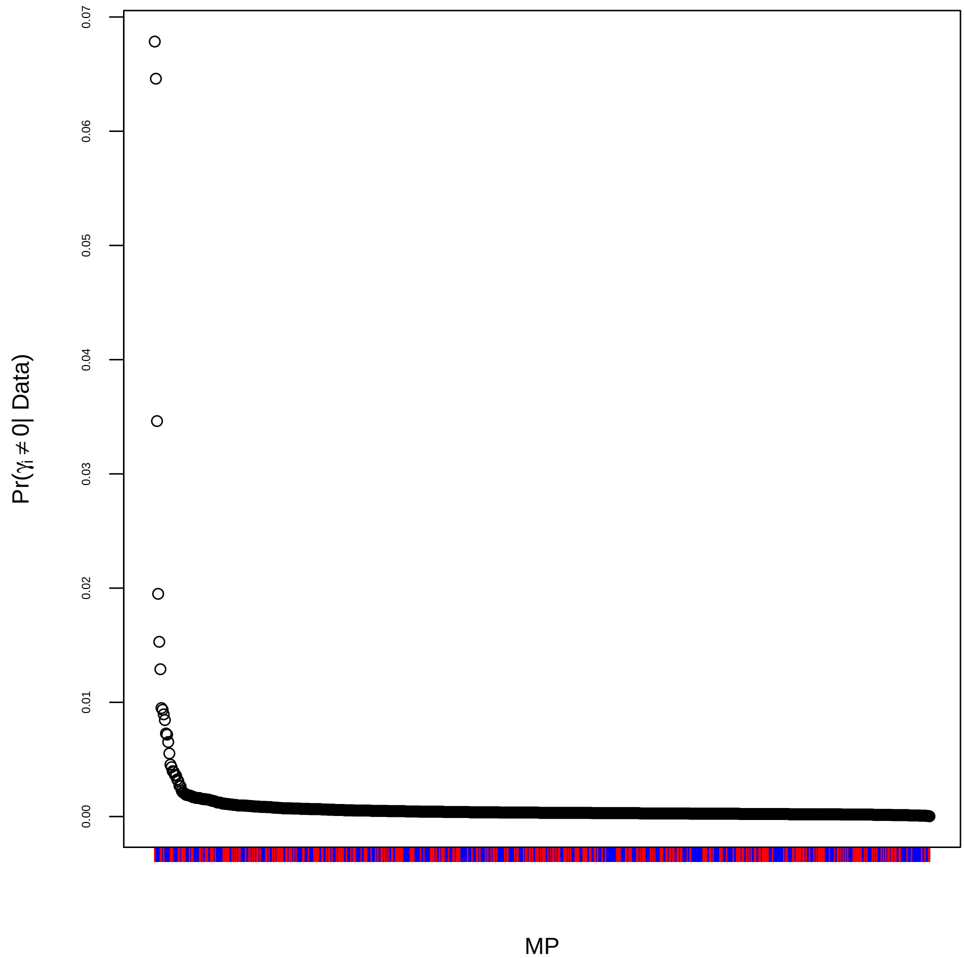


Figure 6: The probability that the abstention pattern of a MP is affected by the position in policy space of a division on electoral reform, 1866-1867. Members of the Liberal party are identified in blue, while members of the Conservative party are identified in red.

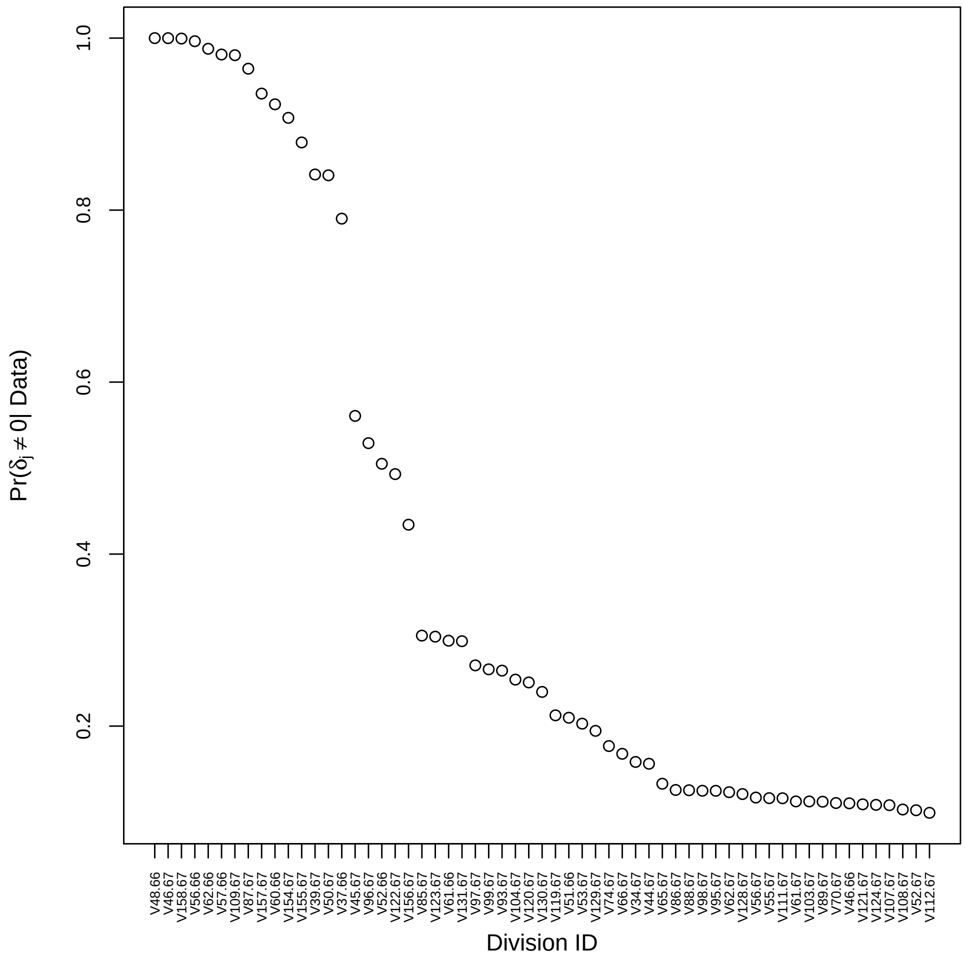


Figure 7: Probability that abstention probabilities of a given division are affected by its position in policy space during the electoral reform deliberatons of 1866-1867.

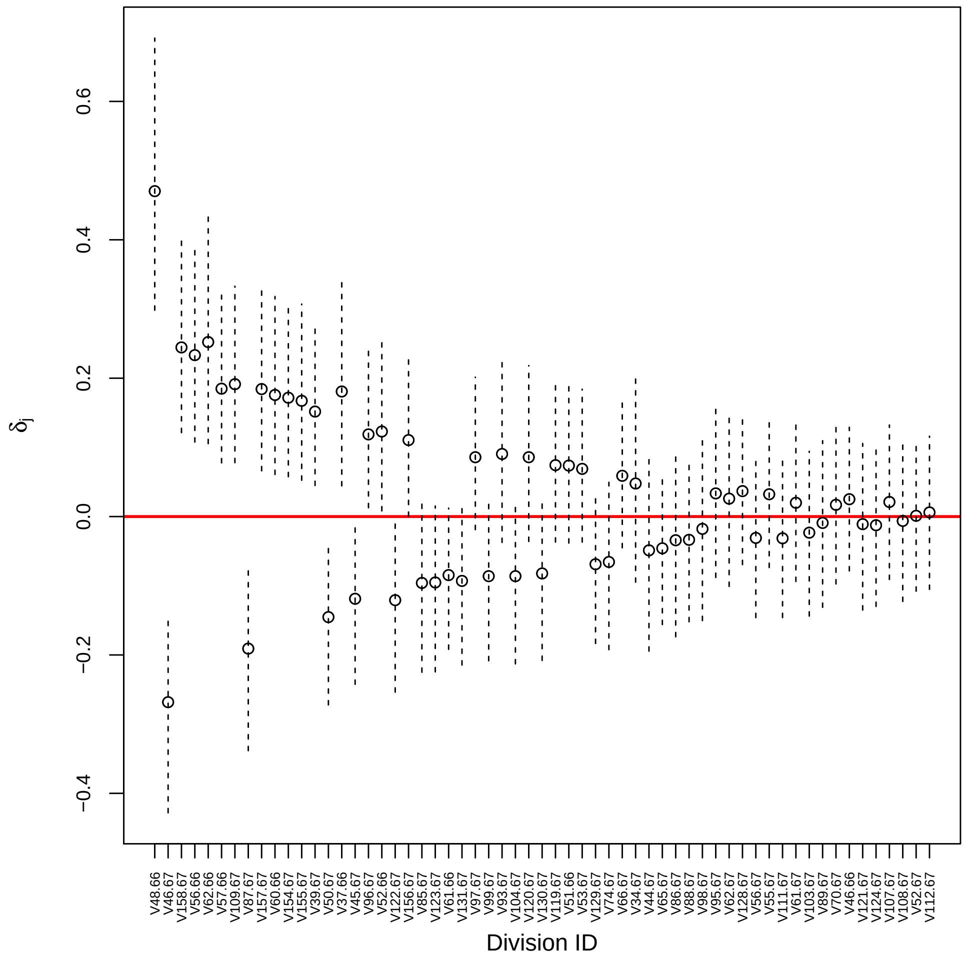


Figure 8: Posterior means and 95% credible intervals for the values of δ_j, conditional on $\delta_j \neq 0$ for divisions taken during electoral reform deliberations, 1866-1867.

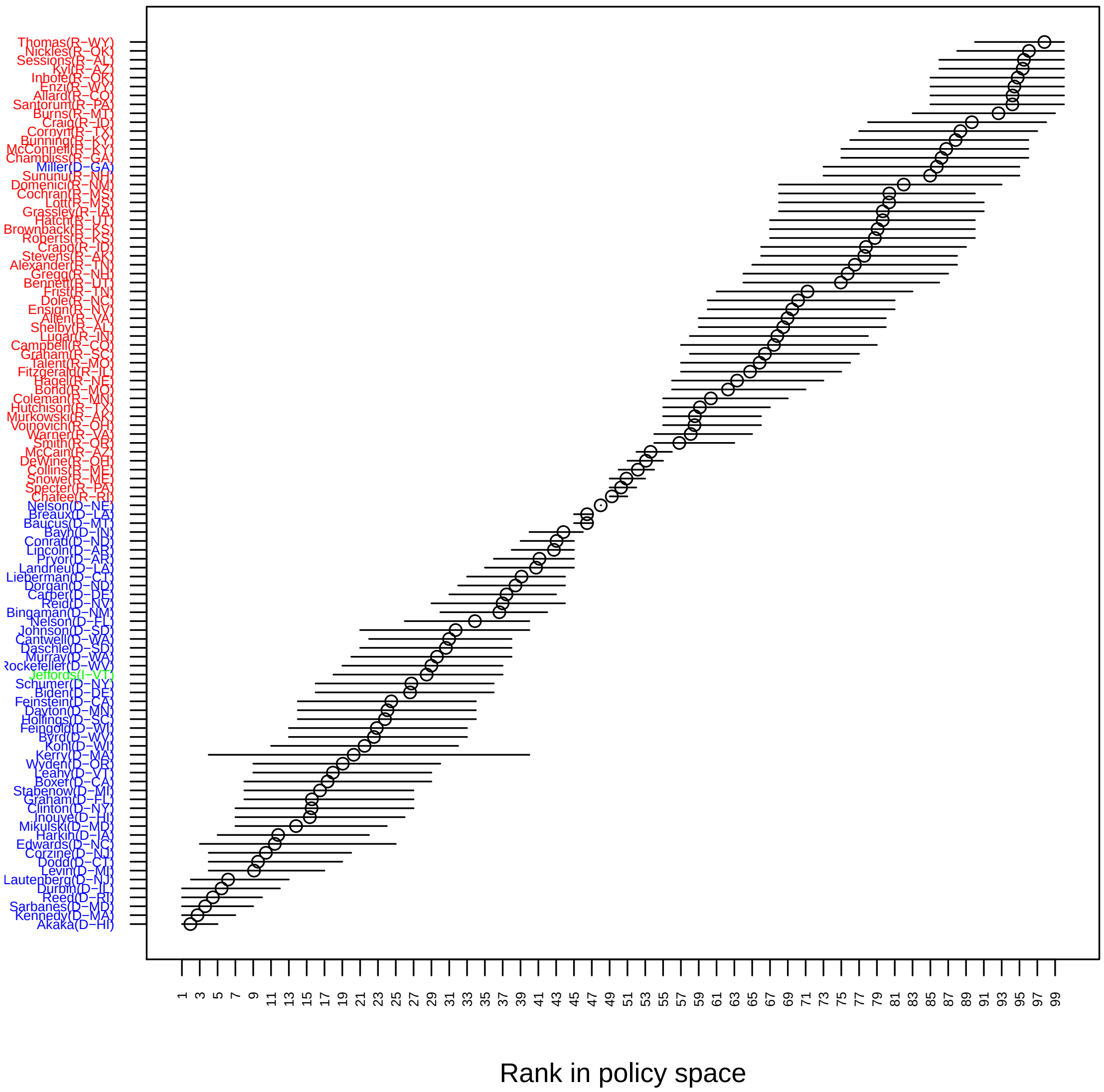
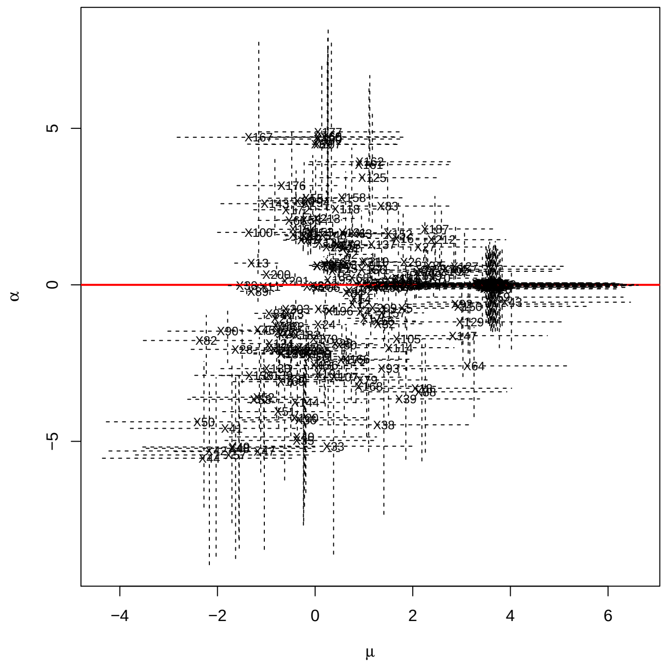
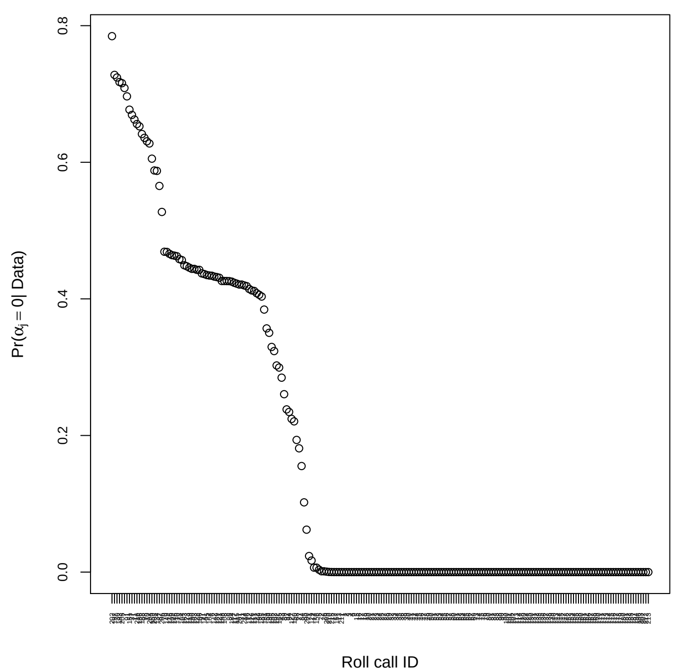


Figure 9: Posterior means and posterior credible intervals for the rank of Senators in policy space in the 108th U.S. Congress. The names of Democrat Senators have been marked in blue, those of Republican Senators in red, and James Jeffords (the only independent) appears in green.

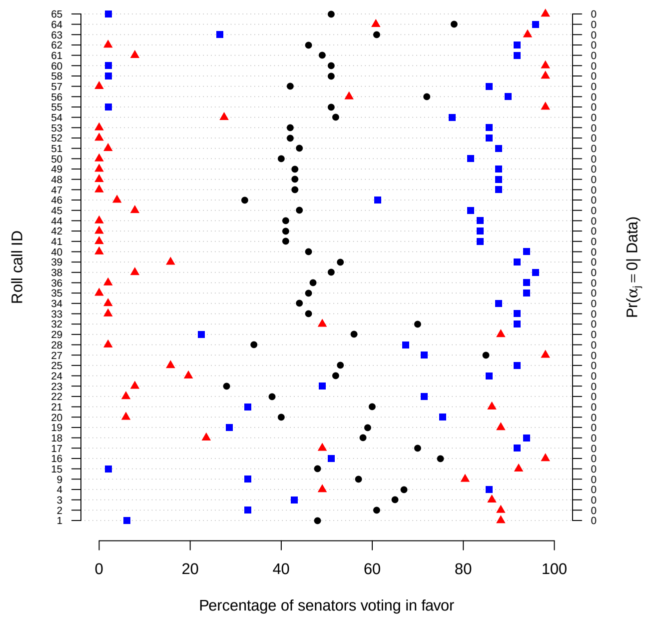


(a)

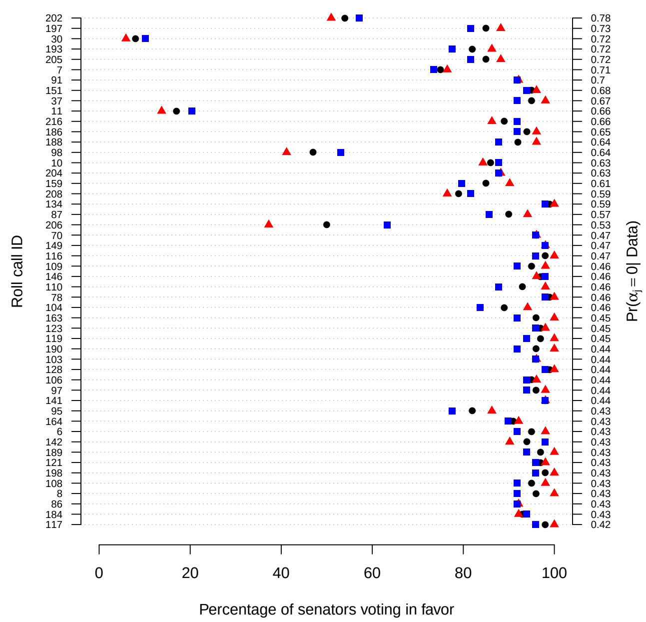


(b)

Figure 10: Panel (a) shows posterior means and posterior credible intervals for bill-specific parameters associated with the voting patterns of the 108th U.S. Congress. Panel (b) shows the posterior probability that a given bill is not informative about the ideal points of the Senators.



(a)



(b)

Figure 11: Percentage of senators voting for the 50 bills with the lowest (top panel) and highest (bottom panel) values of $\mathsf{Pr}(\alpha_j = 0 \mid \text{Data})$. Black circles represent the percentage of the total vote, while blue squares represents the percentage of the republican vote, and red triangles represent the percentage of democrat vote in favor of the bill. Numbers on the right vertical axes correspond to $Pr(\alpha_j = 0 \mid \text{Data})$ for the each of the bills being considered.

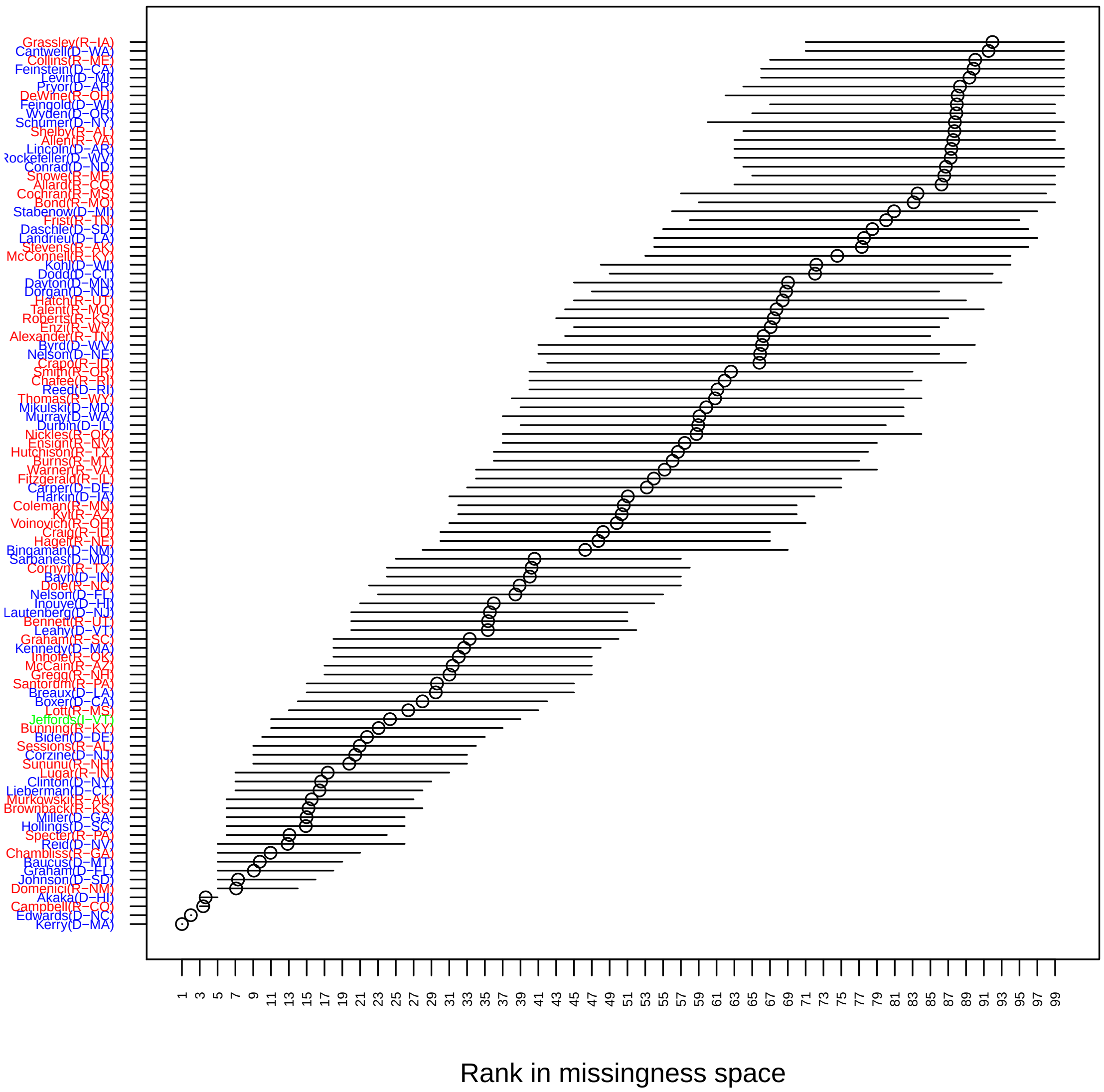


Figure 12: Posterior means and posterior credible intervals for the ranks of Senators in the 108th U.S. Congress in missingness space. The names of Democrat Senators have been marked in blue, those of Republican Senators in red, and James Jeffords (an independent) appears in green.

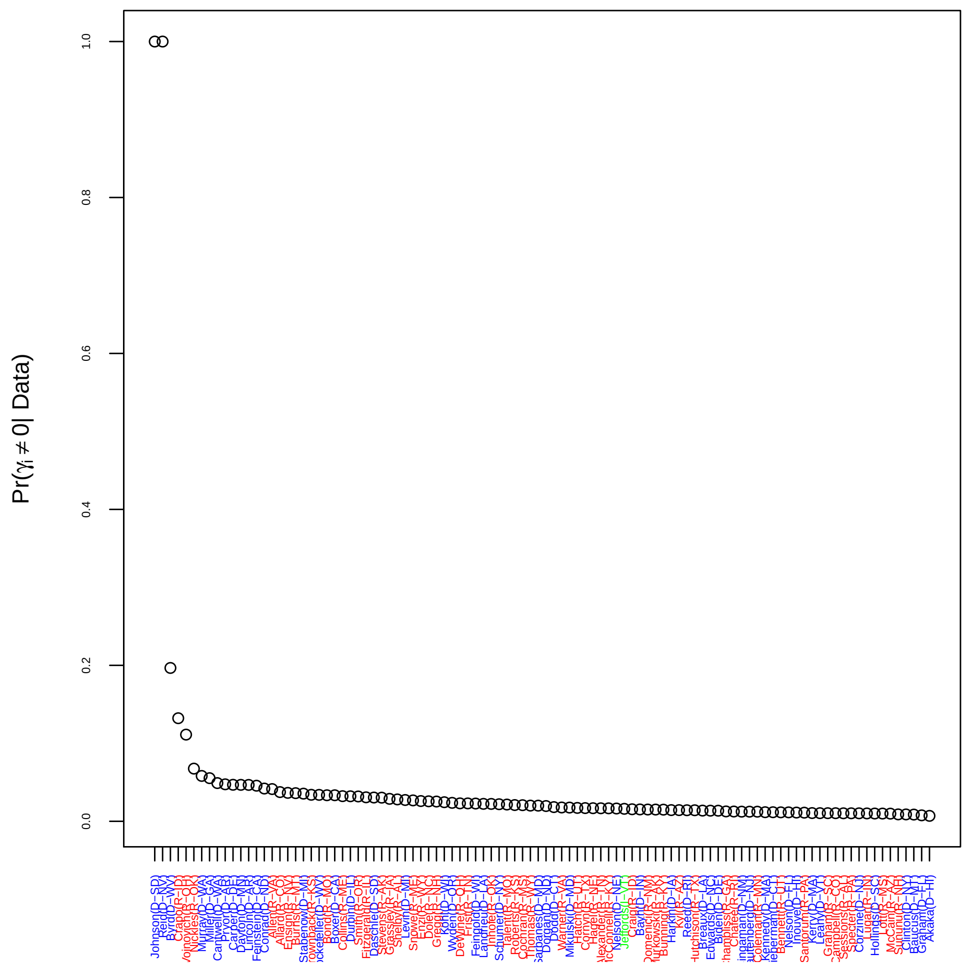


Figure 13: Probability that the voting pattern of a Senator is affected by the location in policy space of a bill in the 108th U.S. Congress. The names of Democrat Senators have been marked in blue, those of Republican Senators in red, and James Jeffords (an independent) appears in green.

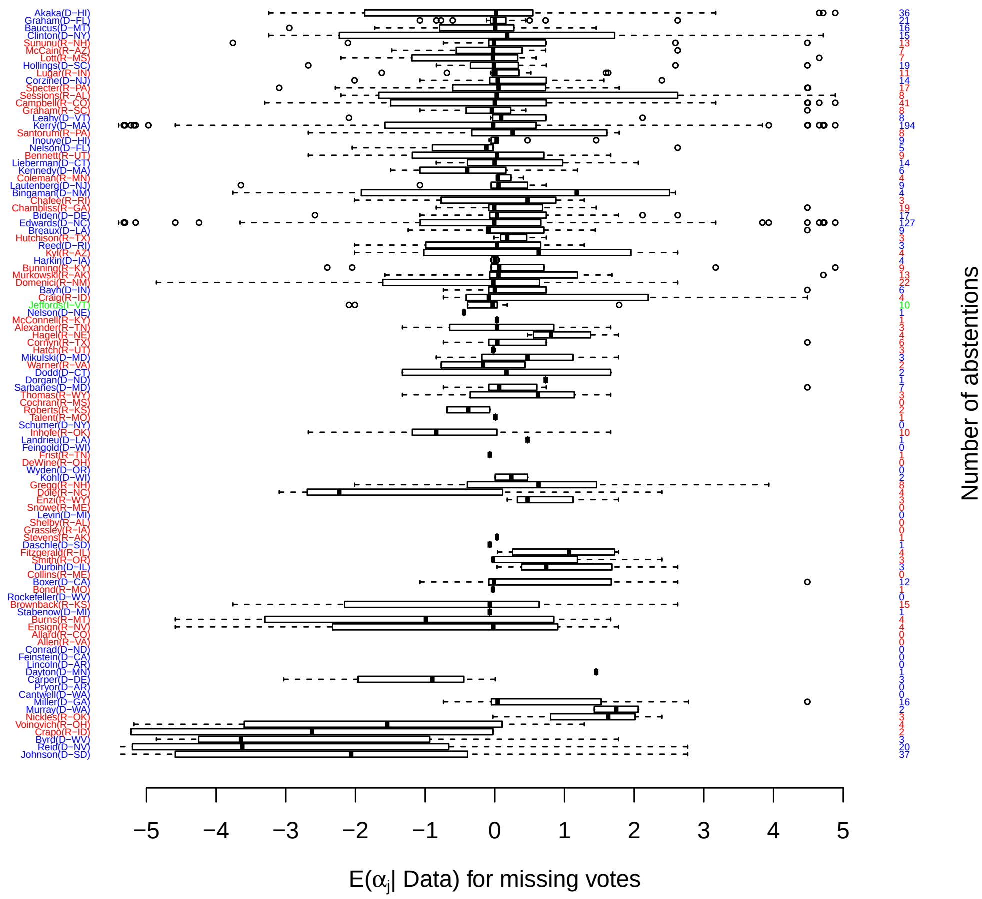


Figure 14: Boxplot of the expected value of the parameters α_j associated with the bills in which each Senator missed a vote. The names of Democrat Senators have been marked in blue, those of Republican Senators in red, and James Jeffords (an independent) appears in green, and they have been ordered according to our estimates of $\Pr(\gamma_i = 0 \mid \text{Data})$. The numbers in the right hand side correspond to the number of abstentions for each Senator.

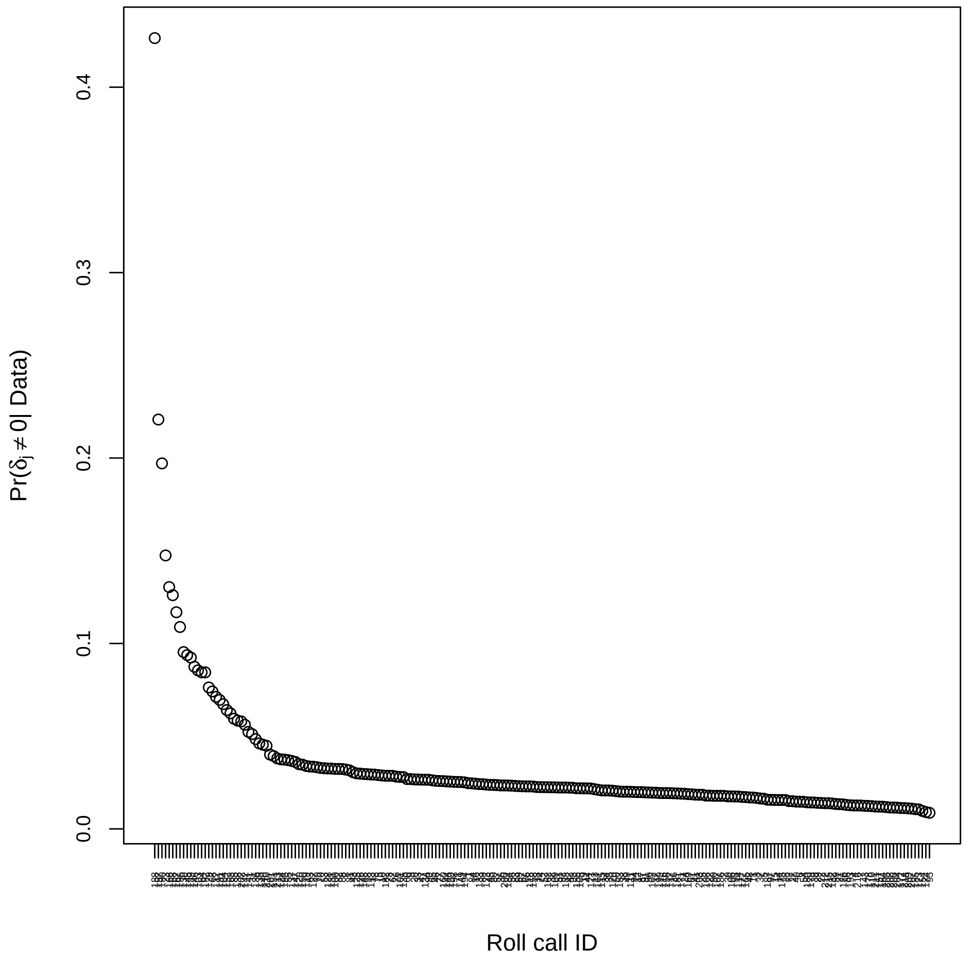


Figure 15: Probability that the voting pattern of a bill is affected by the level of partisanship of a Senator in the 108th U.S. Congress.