

Voting, Speechmaking, and the Dimensions of Conflict in the US Senate*

In Song Kim[†] John Londregan[‡] Marc Ratkovic[§]

August 31, 2015

Abstract

Legislative institutions structure and order the myriad topics addressed by legislators. Jointly considering both roll call votes and floor speech, we show that the contemporary US Senate is ordered around but two dimensions: one ideological and another capturing leadership. We characterize both word and vote choice in terms of the exact same ideal points and policy dimensions. These findings emerge from our method, Sparse Factor Analysis (SFA), designed to combine vote and textual data when estimating ideal points, word affect, and the underlying dimensionality. This contrasts with the single dimension that emerges from an analysis of votes alone, and with the more numerous dimensions that emerge from analyzing speech alone. We then show how SFA can leverage common speech in order to impute missing data, to estimate rank-and-file ideal points using only their words and the vote history of party leaders, and even to scale newspaper editorials.

*We thank Arthur Spirling, Jong Hee Park, Nolan McCarty, Brandon Stewart, Kevin Quinn, Ben Johnson, Tolya Levshin, and Michael Peress for comments on this and an earlier draft. Earlier versions were presented at the 2015 Asian Political Methodology Conference, Political Economy workshop at Centro de Investigacion y Desarrollo de la Educacion Cide and the 2015 Conference on Ideal Point Models.

[†]Assistant Professor, Department of Political Science, Massachusetts Institute of Technology, Cambridge, MA, 02139. Email: insong@mit.EDU, URL: <http://web.mit.edu/insong/www/>

[‡]Professor of Politics and International Affairs, Woodrow Wilson School, Princeton University, Princeton NJ 08544. Phone: 609-258-4854, Email: jbl@princeton.edu

[§]Assistant Professor, Department of Politics, Princeton University, Princeton NJ 08544. Phone: 608-658-9665, Email: ratkovic@princeton.edu, URL: <http://www.princeton.edu/~ratkovic>

1 Introduction

Members of the US Congress cast votes on hundreds of substantive matters every year, touching on issues foreign and domestic, parochial and national. Many scholars argue that legislative voting behavior can be well-summarized by a single dimension of conflict (Poole and Rosenthal, 1997; Krehbiel, 1998; Clinton, Jackman and Rivers, 2004). In this one-dimensional characterization of Congress, each legislator can be identified with a most preferred outcome, or *ideal point*: a single summary measure that conveys the Members' relative preferences across a variety of bills, up to idiosyncratic error. Theorists have long argued that committee and party system serve to reduce the dimensionality of conflict (Shepsle and Weingast, 1981; Cox and McCubbins, 1993; Gilligan and Krehbiel, 1987; Aldrich, 1995), avoiding the chaotic voting behavior identified by McKelvey (1976) and Schofield (1977) while providing useful heuristics to both legislators and voters. From this perspective, the committee and party structure serve to channel a larger political and policy space into this one dimension.

In contrast with this low dimensional interpretation used by both legislative scholars and various formal theorists, recent text-analytic methods spurn parsimony and model word frequency and co-occurrence as generated from scores of latent *topics* (Roberts et al., 2014; Grimmer and Stewart, 2013; Monroe, Colaresi and Quinn, 2008; Blei, Ng and Jordan, 2003). For example, Congressional floor speech has been characterized with forty-two topics and Senate press releases by forty-three (Quinn et al., 2010; Grimmer, 2010). These results introduce a disconnect between the theory underlying vote choice and that underlying word generation. Whereas the spatial model of voting provides a choice theoretic model that creates a close link between policy preferences and voting, no similar model undergirds legislator's choice of words in topic models (Sim, Routledge and Smith, 2014; Lauderdale and Herzog, 2014; Lauderdale and Clark, 2014; Gerrish and Blei, 2012; Slapin and Proksch, 2008). Furthermore, whether scaling votes or modeling word occurrence, the number of underlying dimensions must be assumed by the researcher (but see Hahn, Carvalho and Scott, 2012; Heckman and Snyder Jr., 1997).

This paper offers several advances. Substantively, we find evidence that Congressional institutions coerce a broader policy space into a low-dimensional space over which legislators cast votes. We analyze eight recent Senates, the 105th through the 112th, and we consider both roll call votes and floor speeches.

Two dimensions emerge from our analysis: the first coincides closely with the dimension identified by roll call votes alone, while it offers insights about the ideological connotations attributed by politicians to the terms they utter (Gentzkow and Shapiro, 2010). The second dimension reflects parliamentary organization, with leaders of both parties sharing one end, and the “rank and file” members, including moderates, showing up at the other. When we restrict our attention to votes alone we find the standard one underlying dimension uncovered by existing measures (Clinton, Jackman and Rivers, 2004; Poole and Rosenthal, 1997). Scaling only terms, we find that Congressional speech is well summarized by 5-8 dimensions, with some year-to-year variation. This is consistent with the view that the agenda in the Senate is controlled to reduce the dimensionality of voting, see for example Shepsle and Weingast (1981).

Methodologically we model vote and word choice as a function of the same underlying preference structure for legislators. Our estimator, Sparse Factor Analysis (SFA), encompasses legislators’ preferred outcomes as well as bill and word features. Unlike most existing methods, SFA estimates the dimensionality of the choice space, rather than requiring the researcher to assume it *ex ante*.

By combining terms and votes to estimate the ideological space they jointly occupy, SFA allows us to use politicians’ common vocabulary to estimate most preferred legislative outcomes for non-voting speech makers. To illustrate this potential we reestimate our model excluding first the votes of but ten Senators chosen at random, and then all but the majority and minority leaders and whips. Yet even when we censor the voting records of most of the Senate, the text from floor speech allows us to recover ideal points that correlate with their DW-NOMINATE scores above 0.8. We then use the overlap between the *spectra* of Congressional floor speech and unsigned newspaper editorials of three prominent national newspapers in order to impute most preferred policies for three prominent newspaper editorial boards.

The paper progresses in four parts. First, we set forth the SFA model and discuss its estimation. Second, we present a more informal discussion of assumptions underlying SFA, and we compare it with preexisting methods. Section three is the heart of our analysis, applying SFA to votes and speech from the US Senate. A brief section concludes, with the technical details in a supplemental appendix. Software for implementing the proposed method is publicly available in package `sfa` in the R programming language.

2 SFA: The Proposed Method

In this section, we develop a choice theoretic spatial model that establishes a basic homology between voting and speech: both are anchored to the very same legislator ideal points. Next, we introduce Sparse Factor Analysis (SFA) for estimating these ideal points using vote data, word data, or a combination of the two. The subsequent section discusses some of the key ideas embedded in SFA.

2.1 Voting and Discourse

We now develop a choice theoretic spatial model for both word and vote choice. In this model both the terms employed in speech and votes correspond to “locations” in an underlying ideological space over which legislators have spatial preferences, preferring results that are “closer” to their most preferred outcome. In both settings the location of the legislator’s most preferred outcome is identical, and the dimensionality is the same, tightly connecting the spatial structure of speech giving and voting.

Any model that encompasses voting and speech requires us to keep separate track of votes and terms. To make this easier, we summarize some key elements of our notation in Table 1.

Legislators’ preferences have two salient features—location and dimensionality. We assume that while legislators differ in the ideological locations of their most preferred outcomes, they all have preferences over an underlying ideological space of the same dimensionality. We denote as a_d , $d \in \{1, \dots, D\}$, the weight accorded by all legislators to the d^{th} dimension of policy. Legislators’ preferences depend on the proximity of those outcomes to their most preferred outcomes. Notationally, the most preferred outcome of legislator $l \in \{1, \dots, L\}$ on dimension $d \in \{1, \dots, D\}$ is given by x_l^d . We assume that legislator preferences, word affect, dimension weights, and the dimensions themselves are fixed.

2.2 Modeling Speech

We start with the model for speech. Our notion of legislative speech is one of expression rather than persuasion. Rather than strategically signaling the quality of proposals (Austen-Smith, 1990), we will treat legislative speech as a means of “position-taking,” “credit-claiming,” and connecting to local constituencies (Hill and Hurley, 2002; Maltzman and Sigelman, 1996; Fenno, Jr., 1978, 1977; Mayhew, 1974).

Denote as T_{lw}^* the intensity with which each legislator $l \in \{1, \dots, L\}$ employs term $w \in \{1, \dots, W\}$.

Symbol	Interpretation
Observed Outcomes	
T_{lw}, V_{lp}	Observed term count T_{lw} and vote V_{lp} for legislator l on term w and proposal p
Latent Outcomes	
T_{lw}^*, V_{lp}^*	Intensity with which legislator l uses term w (T_{lw}^*) and supports proposal p (V_{lp}^*)
Model Parameters, Common Across Term and Vote Models	
x_{ld}	Preferred outcome, legislator l and dimension d
a_d	Weight for dimension d
Model Parameters that Differ Across Term and Vote Models	
$\theta_{lw}^{term}, \theta_{lp}^{vote}$	Systematic component of term and vote choice
$g_{wd}^{term}, g_{pd}^{vote}$	Term location (g_{wd}^{term}) and gap between Aye and Nay vote (g_{pd}^{vote}), dimension d
c_l^{term}, c_l^{vote}	Auxiliary legislator parameters
b_w^{term}, b_p^{vote}	Auxiliary proposal parameters
$\tilde{\epsilon}_{lw}^{term}, \tilde{\epsilon}_{lp}^{vote}$	Random component of term and vote choice

Table 1: **Elements of the model.** For clarity, we present the variables central to our analysis. The ideal point, x_{ld} , and dimension weight a_d for each dimension d are common across the word and vote models, providing an explicit link between the two. The remaining parameters and outcomes vary between the word and vote model, though they perform a similar function in each.

Term w has an ideological location g_{wd} in each dimension d .

In selecting how intensely to use a term, the legislator considers both its ideological affect and its pertinence. The contribution of ideological proximity is inversely proportional to the distance between the ideological “location” of the term and a legislator’s most preferred outcome:

$$Ideol_l(T_{lw}^*) = -\frac{1}{2}T_{lw}^* \sum_{d=1}^D a_d(x_{ld} - g_{wd}^{term})^2. \quad (1)$$

The closer a term’s ideological content to the legislator’s most preferred outcome, the more prone she will be to employ the term. Absent any context, legislators would simply utter terms closest to their preferred outcomes *ad infinitum*. They do not do so because the *pertinence* of terms also matters. In the context of Congressional discourse, if health care dominates the agenda for the current session, it is relatively easy to make frequent use of ideologically charged terms entwined with health, such as “clinic”, or “seniors”, while it becomes costly for a legislator to veer off topic to mention her heartfelt position on, say, gun control.

To formalize the interaction of ideological affect and topical pertinence in shaping term choice, we characterize pertinence as increasing in the aptness of the term, s_w , and in the baseline verbosity of the legislator, v_l . Pertinence decreases in the frequency with which the term is used: any phrase, however powerful, becomes hackneyed with overuse.

$$Pert_{lw}(T_{lw}^*) = T_{lw}^* \left(v_l + s_w - \frac{1}{2}T_{lw}^* - \epsilon_{lw}^{term} \right).$$

We define the legislators’ utility as the sum of ideological agreement and pertinence,

$$U_{lw}(T_{lw}^*) = Ideol_l(T_{lw}^*) + Pert_{lw}(T_{lw}^*).$$

Differentiation gives a maximum of the form:

$$T_{lw}^* = v_l - \underbrace{\frac{1}{2} \sum_{d=1}^D a_d x_{ld}^2}_{c_l^{term}} + s_w - \underbrace{\frac{1}{2} \sum_{d=1}^D a_d (g_{wd}^{term})^2}_{b_w^{term}} + \sum_{d=1}^D a_d x_{ld} g_{wd}^{term} - \epsilon_{lw}^{term}$$

Gathering terms, we have legislator l ’s preferred intensity for term w :

$$\begin{aligned}
T_{lw}^* &= c_l^{term} + b_w^{term} + \sum_{d=1}^D a_d x_{ld} g_{wd}^{term} - \tilde{\epsilon}_{lw}^{term} \\
&= \theta_{lw}^{term} - \tilde{\epsilon}_{lw}^{term}.
\end{aligned} \tag{2}$$

2.3 Modeling Votes

We characterize vote choice with a familiar implementation of the spatial model (Clinton, Jackman and Rivers, 2004; Ladha, 1991). If passed, proposal $p \in \{1, \dots, P\}$ has spatial consequence r_{pd} in each dimension d ; failure to pass results in the spatial consequence q_{pd} . Proposals cover a range of outcomes, for the US Senate these include votes on final passage of a bill, votes on whether to recommit a bill to committee, or even votes on whether to honor “...the victims and heroes of the shooting on January 8, 2011 in Tucson, Arizona.”

Each legislator evaluates each proposal in terms of its proximity to her most preferred outcome. She values the positive alternative as:

$$U_l(\{r_p^d\}_{d=1}^D) = -\frac{1}{2} \sum_{d=1}^D a_d (r_{pd} - x_{ld})^2 + \tilde{\xi}_{lp}^{aye}$$

and the negative option as:

$$U_l(\{q_p^d\}_{d=1}^D) = -\frac{1}{2} \sum_{d=1}^D a_d (q_{pd} - x_{ld})^2 + \tilde{\xi}_{lp}^{nay}$$

where $\tilde{\xi}_{lp}$ denote idiosyncratic preference shocks, whose means may differ from zero, for example if there is *status quo* bias.

The intensity of legislator l 's preference for the positive option on proposal p , denoted V_{lp}^* , is given by the difference between voting in the affirmative and in the negative:

$$V_{lp}^* = U_l(\{r_p^d\}_{d=1}^D) - U_l(\{q_p^d\}_{d=1}^D) \tag{3}$$

Substituting and simplifying¹ gives:

¹For a full derivation, see the technical appendix.

$$\begin{aligned}
V_{lp}^* &= c_l^{vote} + b_p^{vote} + \sum_{d=1}^D a_d x_{ld} g_{pd}^{vote} - \tilde{\epsilon}_{lp}^{vote} \\
&= \theta_{lp}^{vote} - \tilde{\epsilon}_{lp}^{vote}
\end{aligned}
\tag{4}$$

where $\tilde{\epsilon}_{lp}^{vote}$ has mean zero:

$$\tilde{\epsilon}_{lp}^{vote} = \tilde{\zeta}_{lp}^{nay} - \tilde{\zeta}_{lp}^{aye} - (E\{\tilde{\zeta}_{lp}^{nay}\} - E\{\tilde{\zeta}_{lp}^{aye}\})$$

The parameters c_l^{vote} and b_p^{vote} are idiosyncratic amalgams of spatial, vote, and voter specific parameters, while $g_{pd}^{vote} = r_{pd} - q_{pd}$ is the gap between the ‘‘Aye’’ and the ‘‘Nay’’ alternatives for proposition p . For each dimension, the parameters x_{ld} and a_d are the same spatial preference parameters the legislators use to measure their affinity for terms when speaking.

2.4 Operationalizing the Model

In estimating SFA, we offer two major advances. First, through a judicious selection of cutpoints, we can connect term counts and votes in the same latent space. Second, through placing a sparsity-inducing prior on each dimension weight, we can recover estimates of how many dimensions are likely non-zero. Much of the model can be estimated using well-established Bayesian methods (e.g., Jackman, 2009), so we emphasize our major contributions. A complete description of estimation is deferred to a technical appendix.

Connecting Terms and Votes

First, we connect the two sets of latent variables, T_{lw}^* and V_{lp}^* to each other. To do so, we assume:

$$\tilde{\epsilon}_{lp}^{vote} \stackrel{\text{i.i.d.}}{\sim} N(0, 1) \text{ and } \tilde{\epsilon}_{lw}^{term} \stackrel{\text{i.i.d.}}{\sim} N(0, 1)$$

This places both intensities on the same latent scale.

Next, we connect the latent term variable, T_{lw}^* , to the observed value, T_{lw} . Define as $\{\tau_c\}$, $c \in \{0, 1, 2, \dots, \infty\}$ the cutpoints that tessellate the real number line, mapping the latent space to the observed space as:

$$T_{lw} = \begin{cases} 0 & \text{iff } -\infty < T_{lw}^* \leq \tau_0 \\ 1 & \text{iff } \tau_0 < T_{lw}^* \leq \tau_1 \\ \dots & \\ k & \text{iff } \tau_{k-1} < T_{lw}^* \leq \tau_k \\ \dots & \end{cases} \quad (5)$$

Likewise, we can connect our vote preference intensities with voting outcomes:

$$V_{lp} = \begin{cases} 0, & -\infty < V_{lp}^* \leq \tau_0^{vote} \\ 1, & \tau_0^{vote} < V_{lp}^* \leq \infty \end{cases} \quad (6)$$

where we make the standard assumption that $\tau_0^{vote} = 0$.

The conditions specified in expression (5) suggests an ordered probit model of McKelvey and Zavoina (1975)(see also Jackman and Trier, 2008). However, given the vast heterogeneity in word counts in a typical Congressional corpus, ranging from zero to several thousand, with many empty count categories, the ordered probit model is not a practical alternative.

We connect the cutpoints with the observed counts, T_{lt} , using the empirical distribution function for term counts:

$$\hat{F}(c) = \frac{1}{LW} \sum_{l=1}^L \sum_{w=1}^W \mathbf{1}(c \leq T_{lw})$$

We then estimate the cutpoints as:

$$\tau_c | \beta_0, \beta_1, \beta_2 = \beta_0 + \beta_1 \hat{F}(c-1)^{\beta_2}$$

assuming that

$$\hat{F}(-1) = 0 \tag{7}$$

$$\beta_0 = \tau_0 = \hat{F}(0)$$

$$\beta_1, \beta_2 > 0$$

This formulation offers three advantages. First, using $\hat{F}(c)$ instead of c guards against outliers (Murray et al., 2013). Second, most term-document matrices contain many zeros, so we model this zero-inflation directly. Our cutpoints ensure that the predicted probability of observing 0 is exactly equal to the actual sample fraction $\hat{F}(0)$. Third, the quasi-linear form allows for a flexible, but still monotonic, set of estimated cutpoints.

Substituting the formulas for the cutpoints into equations (5), we have the usage frequency probabilities for each observed term:

$$Pr\{T_{lw} = n|\cdot\} = \begin{cases} \Phi(\hat{F}(0) - \theta_{lw}^{term}) & \text{if } n = 0 \\ \Phi(\hat{F}(0) + \beta_1 \hat{F}(n)^{\beta_2} - \theta_{lw}^{term}) & \\ -\Phi(\hat{F}(0) + \beta_1 \hat{F}(n-1)^{\beta_2} - \theta_{lw}^{term}) & \text{if } n \geq 1 \end{cases} \tag{8}$$

We can likewise substitute from equation (6) to obtain a similar formulation for the vote data:

$$Pr\{V_{lp} = \tilde{V}_{lp}|\cdot\} = \Phi((2\tilde{V}_{lp} - 1)\theta_{lp}^{vote}) \tag{9}$$

for $\tilde{V}_{lp} \in \{0, 1\}$.

A final concern is the relative weight to give votes and terms when formulating the factors. People tend to use more words than they register votes, with differences often an order of magnitude. We wish to construct ideal points such that word information does not swamp the vote data.

We offer two means of striking a balance between the two sources. The first is to treat the balance as a tuning parameter and conduct several analyses, varying this parameter. This is the approach we follow below. Below, we present results from three models: one weighing votes heavily, one weighing words heavily, and a third where the two are balanced.

For a data-driven means of selecting this balance, we offer an external criterion that lets the data inform what value of word and vote information generates the most discriminating ideal points. Denote as α a parameter for the relative weight between votes and words. We vary α from 10/1, where most of the weight in constructing ideal points is given to votes, to 1/10, where most of the weight is given to words.

Intuitively, we prefer strong dimension weights and ideal points that provide maximal discrimination among individuals. The worst case scenario would be returning dimension weights of zero and identical ideal points for each individual. The more the model recovers strong dimensions and discriminates among individuals, the better. To that end, we select α to minimize the following statistic,²

$$disc(\alpha) = \sum_{d=1}^D \sum_{l=1}^{L_t} \sum_{l'=1}^L \bar{a}_{d,\alpha}^2 (\bar{x}_{ld,\alpha} - \bar{x}_{l',\alpha})^2 \quad (10)$$

with $\bar{x}_{ld,\alpha}$ and $\bar{a}_{d,\alpha}^2$ are the estimated posterior means of $x_{ld,\alpha}$ and $a_{d,\alpha}^2$ for a given value of α . The statistic is zero when there is no discrimination among individuals on ideal points, and grows in the level of discrimination. We do note that the statistic selects a relative ratio of votes to words of about 2:1, with results substantively similar to the 1:1 ratio analyzed below.

Estimating Dimensionality

The legislator, bill, and word means are random effects with common mean, μ , each with a standard normal prior. We apply a more diffuse normal prior with mean zero and variance four to the bill discrimination (g_{pd}) and term level (g_{wd}) parameters and to the preferred outcomes (x_{ld}). These diffuse priors are “uninformative” in the sense that the prior serves mainly to discourage extreme outliers. Specifically, we assume:

$$\begin{aligned} c_l, b_w &\stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, 1) \\ \mu &\sim \mathcal{N}(0, 1) \\ g_{wd}, x_{ld} &\stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 4) \\ \log(\beta_1), \log(\beta_2) &\stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1) \end{aligned} \quad (11)$$

These priors are conjugate with the likelihood. Thus, conditional on β_1, β_2 , and the a_d , we achieve closed form conditional posterior densities for the remaining parameters by employing data augmentation

²We are rather grateful to Brandon Stewart for suggesting this approach.

(Park and Casella, 2008) to implement a Gibbs sampler.

One of our primary goals is to estimate the number of underlying dimensions along with the ideal points and bill/word weights for each. We do so by placing a sparsity-inducing prior over the elements $\{a_d\}_{d=1}^D$. The prior contains a thresholding parameter, λ , such that larger values of λ set more parameters to zero. We estimate λ within a Bayesian framework (Park and Casella, 2008). Specifically, we place a prior over $\{a_d\}_{d=1}^D$ of the form:

$$\Pr(a_d) = \frac{1}{2\lambda} e^{-\lambda|a_d|}$$

We calculate \hat{a}_d^{ML} , the maximum likelihood estimate of a_d , through standard operations on the matrix of latent variables (Golub and van Loan, 1996), while given λ , the posterior mode can be calculated as in (Donoho and Johnstone, 1994):

$$\hat{a}_d^{MAP} = \begin{cases} \hat{a}_d^{ML} - \lambda, & \hat{a}_d^{ML} > \lambda \\ 0, & \hat{a}_d^{ML} \leq \lambda \end{cases}$$

We estimate λ within a Bayesian framework. This turns the dimension selection step into a simple truncation. We then estimate the number of dimensions as the number of dimensions with a non-zero posterior mode.

We close our estimation algorithm using a Hamilton Monte Carlo method (Neal, 2011) to simulate draws from the posteriors of β_1 and β_2 , while we use data augmentation to restore conjugacy to the dimension weights (Park and Casella, 2008). Details are available in an online technical appendix.

3 A Discussion of Our Method

Political science has been at the forefront of innovative methods for analyzing text and for understanding votes, offering both cutting edge work in text analysis and a rich history of scaling vote data. SFA fits into this tradition, while offering several advantages over existing methods. For a complete review of text analytic methods, we refer the reader to Grimmer and Stewart (2013).

While our discussion has focused on situations in which all legislators provide both votes and speech,

there are situations in which we observe the terms spoken by legislators, but we either lack their votes, or, as in the case of “whipped” votes or institutional norms towards conformity, we doubt their sincerity (Kellerman, 2012; Epstein, Segal and Spaeth, 2001). In such situations we can use legislative speeches to infer the preferences that would have guided legislators’ votes had they been able to cast them freely.

Second, SFA estimates the ideological location of term. Existing studies estimate the political affect of terms by attributing left leaning content to those used by Democrats, and rightward import to those spoken by Republicans (Gentzkow and Shapiro, 2010), or modelers posit that terms and votes arise from two conditionally independent data generating processes (Lauderdale and Herzog, 2014; Gerrish and Blei, 2012). SFA scales terms and votes simultaneously, providing natural structural estimates of word affect.

Topic models have grown increasingly popular in recent years (e.g., Roberts et al., 2014; Quinn et al., 2010; Grimmer, 2010). The primary difference between topic models and SFA is the same as that between cluster analysis and factor analysis (Cattell, 1978). Topic models estimate clusters of word that co-occur but do not place them in an underlying latent space. Topic models serve as excellent descriptive tools (e.g., Blaydes, Grimmer and McQueen, 2014). If, on the other hand, the researcher seeks to identify the underlying structure of the source of the observed behavior, whether it is speech or voting, then SFA provides a clearer picture. For a more complete discussion comparing SFA to topic models, see the Technical Appendix.

Several works have dealt with mixed response factor analysis, a method for uncovering underlying latent factors when the data are a combination of categorical, continuous, and count data (Mohamed, Heller and Ghahramani, 2008; Collins, Dasgupta and Schapire, 2002). Quinn (2004) converts observed continuous data to a z -scale, and then combines it with a latent z -scale underlying ordered and categorical data. Murray et al. (2013) extends Quinn’s method, through using the inverse empirical CDF to place any class of variable on an underlying z -scale. Like SFA, both are examples of Gaussian copula models, which convert variables to the same underlying scale and then analyze their correlational structure on the transformed scale. SFA differs in that we estimate the underlying dimensionality. Our method is more powerful than that of Murray *et al.*, as we estimate the cutpoints rather than invert the empirical marginal CDF of each variable. We are not the first to estimate underlying ideological dimensionality. Heckman and Snyder Jr. (1997) recast the problem as one of testing nested hypotheses, while Hahn, Carvalho and

Scott (2012) use a spike-and-slab prior over dimension weights. Aldrich, Montgomery and Sparks (2014) show that sufficiently large cross-party variance can mask important within-party dimensions. We differ from these works in combining both vote (binary) and word (count) data.

Slapin and Proksch (2008) offer a means of scaling terms, *Wordfish*. Our model leads to an identical formulation for the systematic component of word choice (θ_{ij}^{term}), albeit that our model has a more formal choice theoretic foundation. Slapin and Proksch exponentiate the systematic value in order to model a non-negative count (see also Bonica, 2014), whereas we convert both to a z -scale. Our approach allows us to place the terms and votes on a common scale. Moreover, SFA, unlike *Wordfish*, estimates the underlying dimensionality.

The original topic model of Blei, Ng and Jordan (2003) has given rise to a flurry of innovation, including several authors who combine terms and votes: Gerrish and Blei (2012), Lauderdale and Herzog (2014) Lauderdale and Clark (2014), and Wang et al. (2013). These works differ from ours in that the estimation of most preferred points and terms are done separately: topics are estimated, and then these weights are used to estimate preferred outcomes for each topic. One approach is to assume topics are independent of votes, producing a two step procedure where topics are estimated and then preferred outcomes are estimated conditional on the resulting topic weights (Lauderdale and Herzog, 2014; Lauderdale and Clark, 2014). Terms and votes may inform each other as in the remaining works, but the relation between preferred outcomes and word choice is implicit and indirect.

SFA differs by offering a tight, and formal, coupling of the ideological preferences that generate terms and votes. Wang et al. (2013, p. 239) argue that political scientists have avoided text models “because of a desire to stay close to the ‘rational actor’ nature of such traditional models.” We concur, and bridge this gap: we offer a model for text generation that is fully compatible with the standard and accepted models of vote choice. Rather than estimating preferred outcomes that vary from topic to topic, we estimate the preferences most likely to have simultaneously produced both the vote and the roll call data.

To illustrate the use and efficacy of SFA, we analyze text and roll call votes from the contemporary US Senate.

4 Results for the Contemporary US Senate

This section consists of two parts. In the first, we apply SFA to votes and floor speeches from the recent US Senate. Combining words with votes, we uncover a stable two dimensional model, with one dimension oriented around ideological matters and the other reflecting Senate leadership. For completeness, we reanalyze our data using only votes, to recover the single dimension familiar to Congress scholars. We also analyze our data using only speech, recovering between five and eight underlying dimensions, depending on the Senate session. In the second part of this section, we illustrate the ability of SFA to extract ideological information from text, even in situations with severe missing data problems.

4.1 The US Senate: 1997-2012

We apply SFA to the last eight sessions of the US Senate. We scale using both votes and words, returning both ideology estimates and our calibration of the underlying dimensionality. Our data come from two sources. Rollcall data come from VoteView.³ For the text data, we rely on floor speeches as gathered by the Sunlight Foundation.⁴ Following standard practice (e.g., Quinn et al., 2010), we stem (eliminating suffixes, so “expenditure” and “expenditures” collapse into “expenditur”) and eliminate stopwords (“of”, “and”, “the”, etc.). In order to account for frequently paired words, such as “price control” or “health insurance,” we consider both one- and two-word sets of words (unigrams and bigrams). We refer to these unigrams and bigrams, after stemming and stop word removal, as “terms.” We generate a legislator-term matrix that contains the number of times each Senator uses each term, pooling all floor speeches across each session. In order to eliminate anomalous words of sporadic frequency such as “anomalous” and “sporadic”, we only maintain terms used by at least ten Senators at least ten times over the course of the Session. Removing these idiosyncratic terms reduces the proportion of zeros in the legislator-term matrix from over 94% to under 20%. A complete summary of number of legislators, votes, and unique terms can be found in the supplemental materials.

³ <http://www.voteview.com/>. Last accessed October 27, 2014.

⁴ <http://www.capitolwords.org/>. Last accessed October 24, 2014. Code used to stem and eliminate stopwords available upon request.

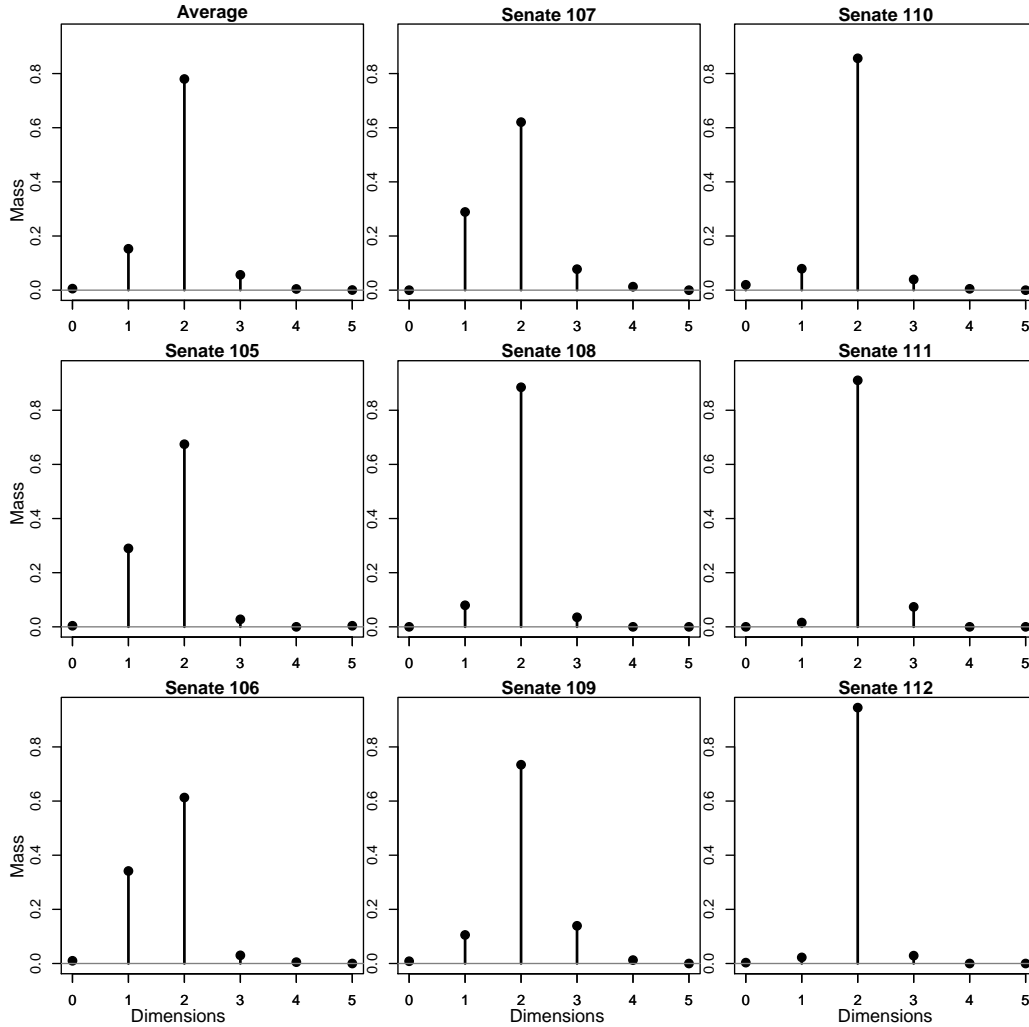


Figure 1: **Posterior density over number of underlying dimensions for the joint word and vote model.** We find a pronounced mode at two dimensions consistently across Senates. The average across all Senates appears in the top left corner.

4.1.1 The Two Dimensions of Vote and Voice on the Senate Floor

Combining word counts for each Senator along with their roll call data, our estimator returns the posterior density over the total number of estimated dimensions, see Figure 1. The average density over the number of dimension parameters merging all Senates is in the top left corner, while the successive sessions are depicted from top to bottom and from left to right. A pronounced mode at two dimensions reappears consistently across Senates.

Not only is the finding of two dimensions consistent, but the two dimensions themselves are stable across sessions. The first closely coincides with the standard ideology dimension uncovered from scaling

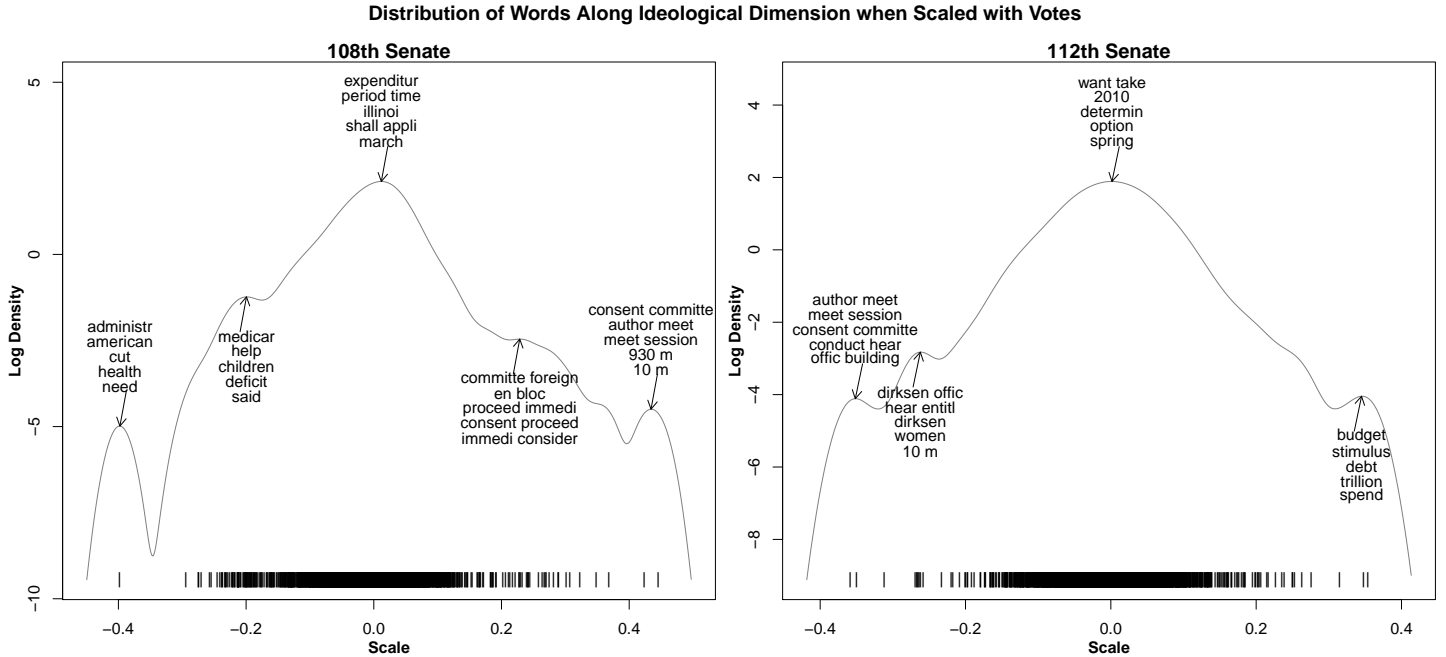


Figure 2: **Log density of term weights, after scaling votes and terms together.** Each local mode is labeled by the five terms closest to that mode. The left figure presents results from the Republican controlled 108th Senate, the right figure contains results from the Democratic led 112th Senate.

roll call votes. The second appears to be a leadership dimension, with party leaders at one end while a variegated mix of “rank and file” partisans and ideological moderates populate the other.

Figure 2 presents the log density of term weights, after scaling votes and terms together. The weights are oriented such that terms more likely to be spoken by Republicans are to the right. Each local mode is labeled by the five terms closest to that mode. The left figure contains results from the 108th Senate, a Republican-led session during President George W. Bush’s tenure. The right figure contains results from the 112th Senate, a Democratic-led session during President Barack Obama’s time as President.

We find a consistent pattern: for the majority party, the most extreme terms relate to parliamentary control words (“consent committee,” “author meet,” “meet session”). For the minority party, the first dimension identifies ideologically relevant terms. For the Democrats during the 108th Senate, these terms included “administr,” as the Democrats soured on the current Presidential administration, and “health,” a centerpiece of the Democratic policy agenda. In the 112th Senate, with the Democrats in the majority, parliamentary control terms switched their ideological polarity, aligning with the Democrats (“author meet,” “meet session,” “consent committee”). The Republican end of this first dimension reflects that

Dimensions Identified by SFA

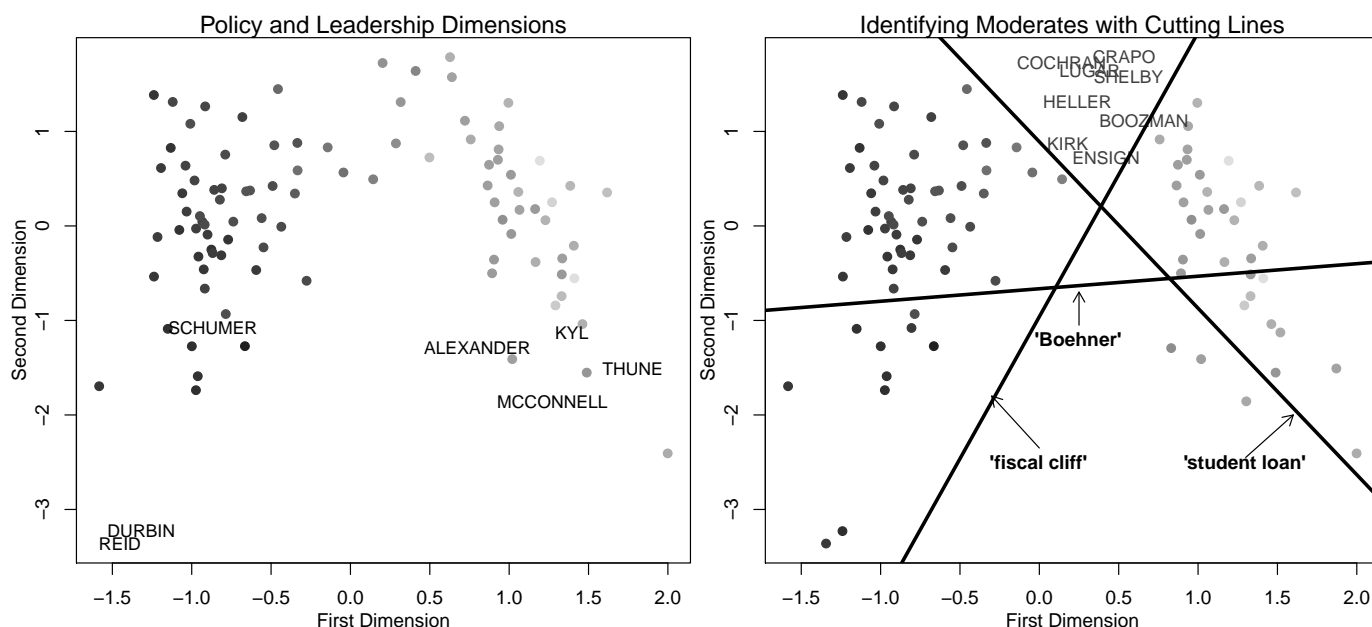


Figure 3: **Latent dimensions estimated by SFA, 112th Senate.** Legislators’ preferred outcomes on the first dimension (x -axis) and the second (y -axis). The left plot labels party leaders, whips, and top chairmen. In the right plot cutting lines separate frequent from infrequent users, of the terms: “Boehner,” “student loan,” and “fiscal cliff.”

party’s programmatic concerns over fiscal balance (“budget,” “stimulus,” “debt,” “trillion”).

Next, we look at the preferred outcomes of legislators from the 112th. Points in Figure 3 are shaded in proportion to their first dimensional DW-NOMINATE score, showing the agreement between SFA and DW-NOMINATE on the first dimension ($\hat{\rho} \approx 0.95$). The left plot labels party leaders, whips, and top chairmen, showing the close relationship between locations on the second dimension and leadership. The first dimension captures the political battle lines, reflecting legislators left *vs* right policy differences, while the second, vertical, dimension reflects differences in the terms selected by leaders versus the rank and file members.

The right plot of Figure 3 contains cutting lines for three terms: “Boehner,” “student loan,” and “fiscal cliff.” The lines were constructed such that legislators on one side are expected to use the phrase above the median number of its raw usage, and on the other side legislators are expected to use the word below its median number of times. We find leaders are more likely to use the word “Boehner,” the House Speaker during this session. Republicans were more likely to use the term “fiscal cliff,” with leaders the most likely. Democrats were more likely to utter the phrase “student loan,” again with leaders the most likely

	SFA Joint Dimension 1	SFA Joint, Dimension 2	SFA Vote, Dimension 1	DW Nom, Dimension 1
SFA Joint (Dim. 2)	0.00 [0.00, 0.01]			
SFA Vote (Dim. 1)	0.85 [0.37 , 0.97]	0.37 [0.16 , 0.9]		
DW Nom. (Dim. 1)	0.83 [0.34 , 0.94]	0.36 [0.16 , 0.89]	0.97 [0.95 , 0.98]	
DW Nom. (Dim. 2)	0.27 [0.16 , 0.35]	0.12 [0.05 , 0.29]	0.34 [0.24 , 0.42]	0.28 [0.19 , 0.33]

Table 2: **Correlation between Estimates, by Model.** The correlation between first- and second-dimension estimates for the SFA model using terms and votes (*SFA Joint*) and for DW-NOMINATE, plus correlations with the first dimension SFA “votes only” estimates (*SFA Vote*). Each entry contains the mean correlation across years with the range across Senates below.

to employ the term. SFA identifies a group of Republican moderates in the top “V”, here we label them by name. These moderates are not likely to use either “student loan” or “fiscal cliff.”

4.1.2 Votes Only

For the sake of comparability with other analyses, we set aside our text data and scale the roll call votes by themselves. We find the same single dominant dimension identified by other analysts (Diaconis, Goel and Holmes, 2008; Clinton, Jackman and Rivers, 2004; Poole and Rosenthal, 1997), for example the scores from SFA and the widely used DW-NOMINATE estimates correlate well above 0.95, even higher than the correlation between DW-NOMINATE and the first dimension of the SFA joint model, with Democrats appearing on one side, Republicans on the other, and more extreme members of each party towards the periphery. While it has become almost a standard practice for analysts to impose *a priori* a single dimension on voting in recent sessions of Congress (but see Aldrich, Montgomery and Sparks, 2014), SFA allows us to estimate the dimensionality of voting for each session. For each session, we estimate with a posterior mass of 100% that the true underlying vote model is one-dimensional.

How do the dimensions encountered by SFA compare with those estimated by other analysts? The correlations among estimates across several different models appear in Table 2. Each entry contains the mean correlation across years, with the range across years reported below the average. The first dimension NOMINATE and first dimension of SFA that jointly uses data on both votes and speech correlate above 0.85 for every Senate except the 107th, where majority control shifted due to Senator James Jeffords’s departure from the Republican Party. While they are not measuring exactly the same behavior, they are

very similar. In contrast, the second dimension of the joint SFA model captures a markedly distinct facet of the data, having to do with the exercise of parliamentary leadership, than does the second dimension of NOMINATE.

4.1.3 Speech Only

We also apply SFA to the terms in isolation from the votes. This is not our preferred model, as it ignores vote data, yet SFA still uncovers structure in the text data. The posterior density of estimated dimensionality for pooled floor speeches can be found in Figure 4. Results across all sessions are in the top left corner while the remaining sessions follow in order from top to bottom and from left to right. In contrast with the high concentration of probability on two dimensions in our preferred model, when we exclude the valuable information contained in votes and analyze oratory alone, we obtain a somewhat more diffuse density that accords a 75% probability to there being between five and eight dimensions, and a probability of over 95% that the underlying dimensionality is within the range [4, 11]. Looking at individual sessions, we find a similar dimensionality, albeit with some year-to-year variation.

To provide a better idea of the content of these dimensions of speech, Figure 5 contains the top ten words at each of the first six dimensions of the 112th Senate.⁵ We note that the positive and negative level distinction along the y -axis is wholly arbitrary, as we only identify term levels up to a sign. Looking at the first column, we find that the first dimension starts with a set of non-controversial terms. These include parliamentary procedural terms (as opposed to parliamentary control terms) such as *today wish*, *madam rise*, and *colleague support*. Also on the non-controversial side are martial terms with universally positive affect during this Congress such as *army*, *air forc*, and *deploy*. On the other side are words that will be used in to differentiate issues in other dimensions, such as *tax*, *vote*, and *peopl*. The other dimensions have at their extremes words connoting some underlying dimension of policy. For example, the second dimension ranges from judiciary and women’s issues at one end to fiscal concerns at the other; the fourth goes from a broad set of social welfare concerns to the consideration of judicial nominees. These lower dimensions adapt to the issues of the day. Tobacco, for example is present in the 105th Senate; Iraq comes and goes as an issue, and health care goes from dealing with seniors and Medicare in the 107th Senate to dealing with students and families in the 112th.

⁵ Results from the 105th–111th Senates are available with the supplemental materials.

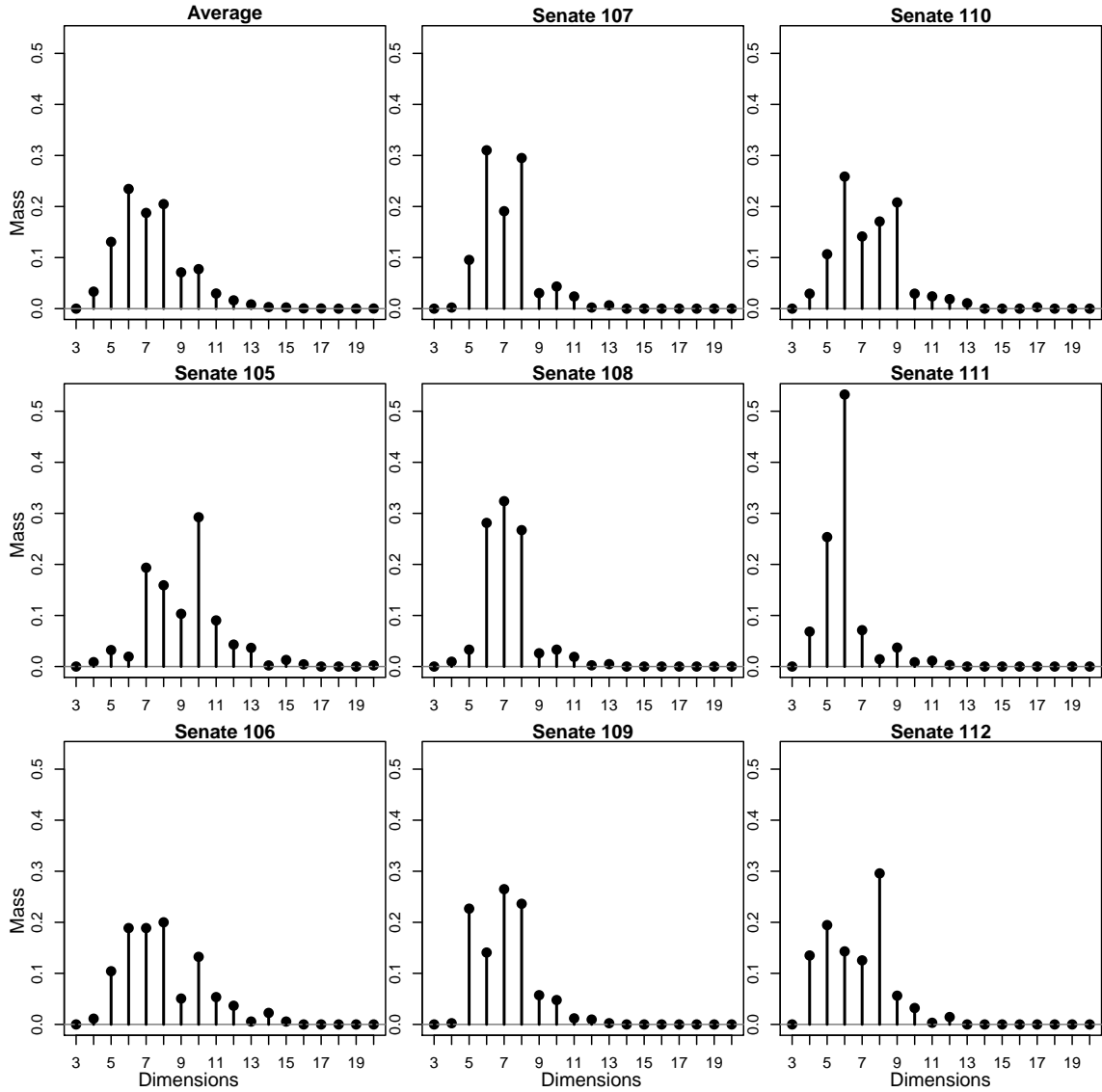


Figure 4: **Estimated underlying dimensionality for Senate floor speeches.** Results across all sessions are in the top left corner and remaining sessions follow.

While scaling terms alone is not our preferred model, the exercise does suggest the utility of SFA for text analytics. An informative depiction of US Senate speech emerges, with a procedural dimension plus several others that chart the shifting substance considered by the Senate. Even without including votes in our analysis, SFA selects a relatively parsimonious representation of the Senate as compared with those found in earlier studies (e.g., Quinn et al., 2010).

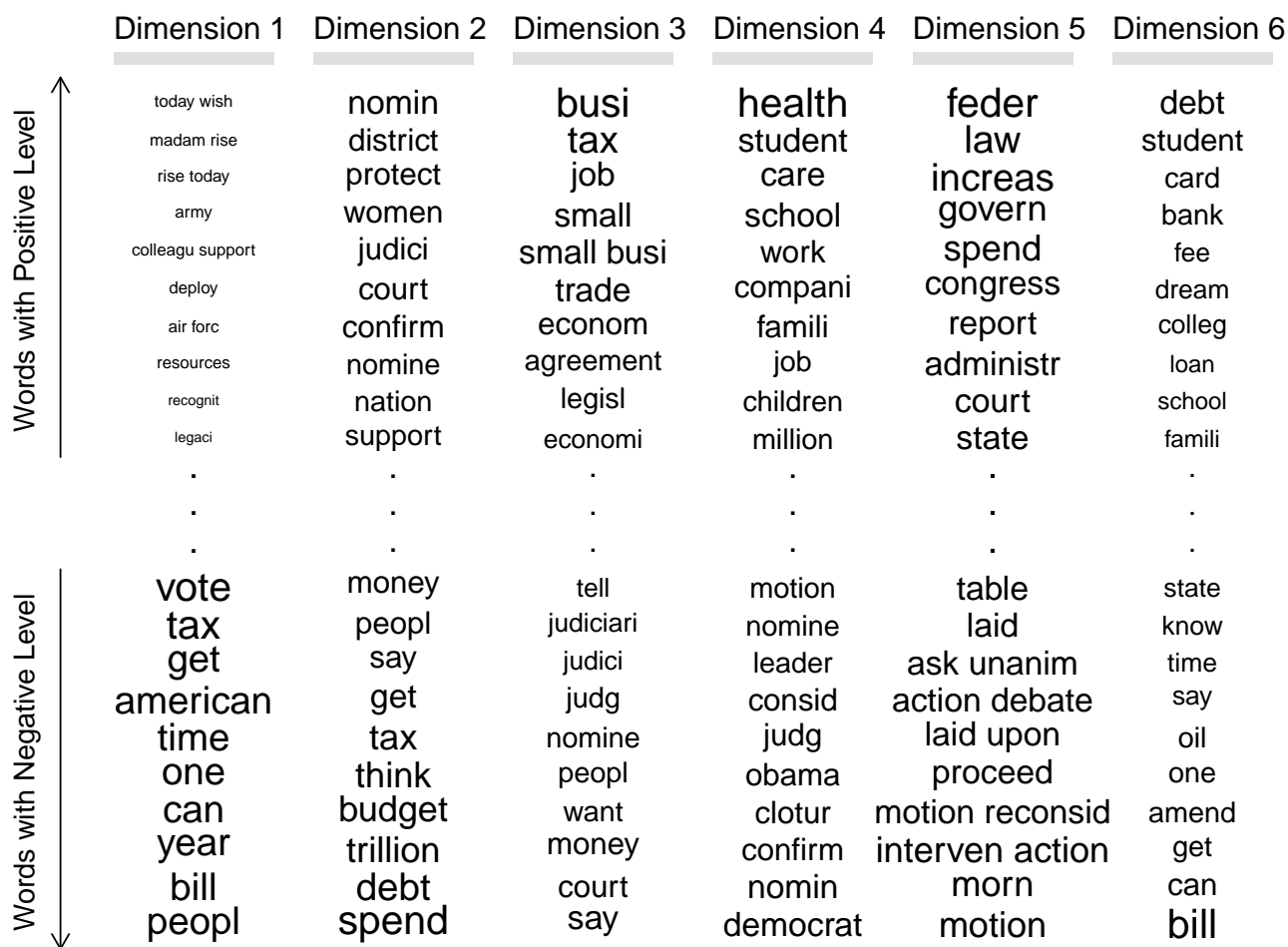


Figure 5: **Extreme Terms by Dimension, 112th Senate.** Extreme terms for the first six dimensions as estimated by SFA from the 112th Senate. The typesize of each term is proportional to the absolute value of the associated coefficient; terms earning positive coefficients appear in the upper part of the panel, those assigned negative coefficients are presented in the lower segment.

4.2 Bridging Votes with Words

SFA offers the possibility of exploiting the link between political speech by individuals who never vote with the prose uttered by those who do. To explore this potential we undertake a pair of exercises using data from the 112th Senate, for which we already possess clear estimates of legislators’ preferred outcomes. We explore how well we could have recovered these results if some of the legislators’ votes had been censored, forcing us to impute the preferences of those legislators based only on their floor speeches. Next, we merge our data on votes and speech in the 112th Senate with the language contained in contemporary unsigned newspaper editorials from the *New York Times*, the *Wall Street Journal*, and the *Washington Post* to

Estimating Ideology when Ten Legislators' Vote Data is Missing Completely at Random

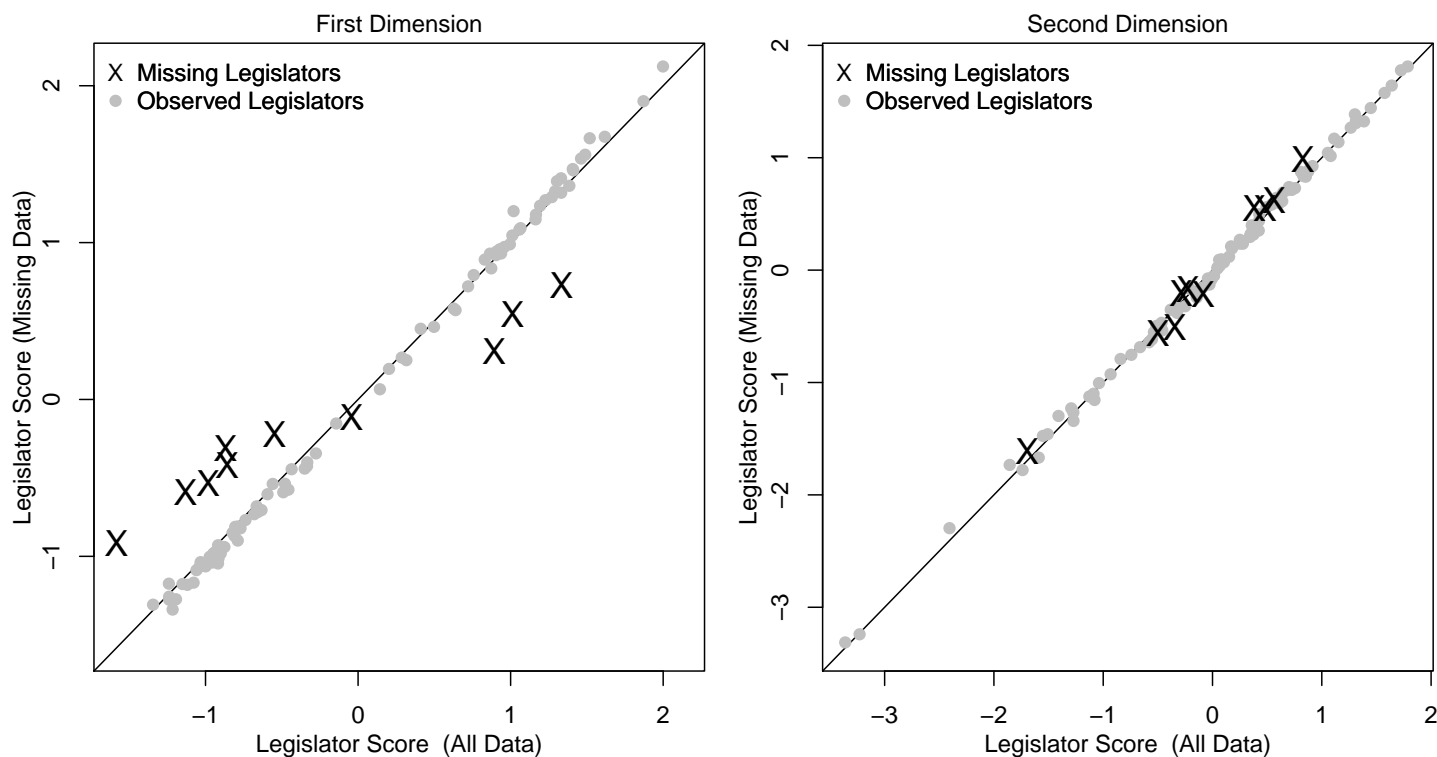


Figure 6: **Estimated Ideal Points for Ten Legislators Missing at Random.** The lefthand panel compares the censored and uncensored estimates (marked by X 's) of the preferred outcomes for the ten randomly censored legislators on the first dimension, while the righthand panel makes the analogous comparison for dimension two.

impute the ideological preferences of their respective editorial boards.

First, we randomly discard the votes cast by ten legislators selected completely at random, coding *all* of their votes as “missing,” while we maintain all of their speech data. The left and right panels of Figure 6 plot the imputed versus fitted values (“ X ”) for the dropped legislators, for the first (left) and second (right) dimension. SFA recovers reliable first-dimension preferred outcomes well, except for some expected attenuation bias. The second dimension ideal points are recovered almost exactly. We remind the reader that the first SFA dimension coincides closely with the dimension that emerges from an analysis of the votes alone, and so we might expect it to be more affected by the loss of voting data, while the accuracy of our second dimension estimates, which are dominated by speech data, would be expected to suffer less from the censorship of the votes.

We next offer a more challenging test of SFA’s abilities. For this analysis, we coded all vote data except

Estimating Ideology when Only Leaders' Votes are Informative

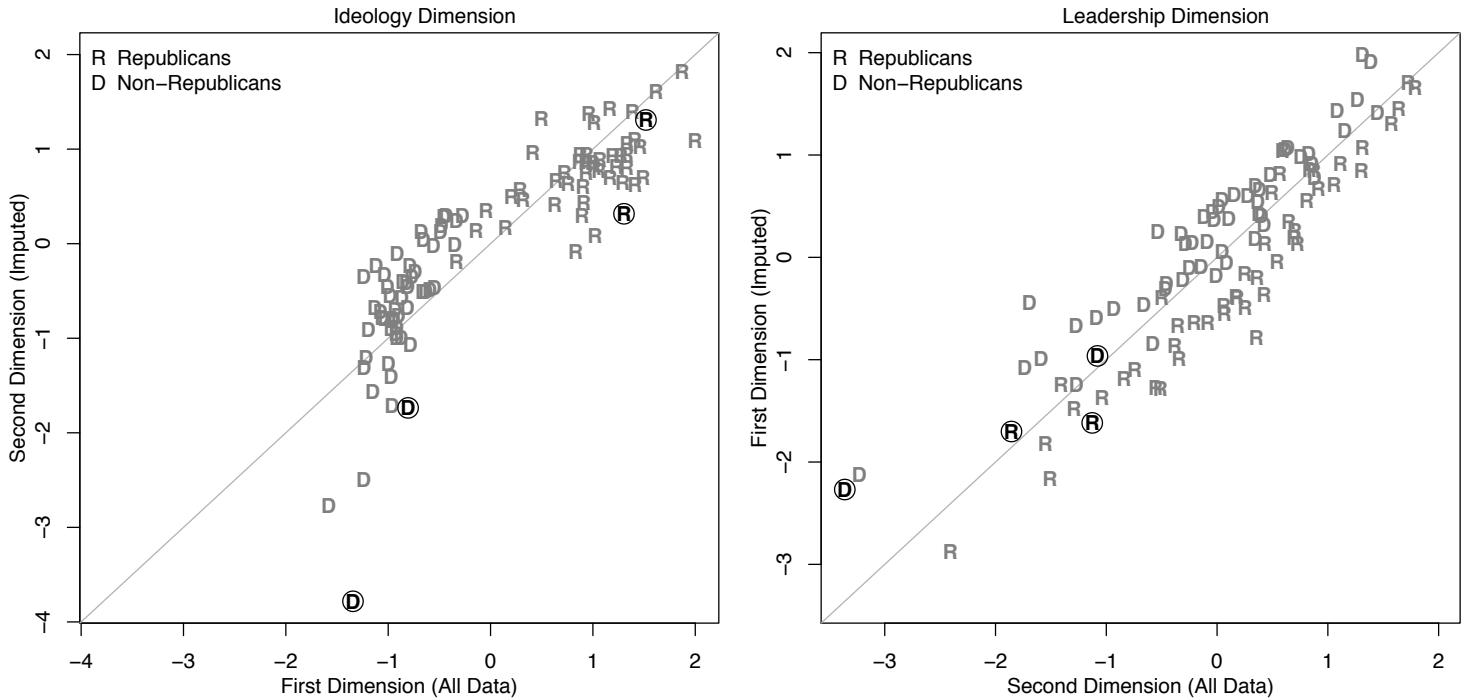


Figure 7: **Estimated Ideology when Only Leaders' Votes are Informative.** The voting dimension estimates appear in the left panel, with the censored estimates measured on the vertical (y -axis) while the uncensored ones appear on the horizontal (x -axis). In the censored data the salience of the voting dimension drops, so that it becomes the second dimension. The righthand panel exhibits the leadership dimension, again the censored estimates correspond with the vertical (y -axis) and the uncensored ones coincide with the horizontal (x -axis).

for the party leaders and whips as missing, while maintaining all speech data. This left a vote record for less than 4% of the Senate. We then compared the SFA ideal point estimates to the SFA estimates using everyone's speech, but only leaders' votes. Essentially our exercise in censorship diminishes the importance of the dimension related to voting. When we estimate the censored data we again recover two dimensions, but their order is reversed, with the voting dimension becoming noisier, and falling into second place, while the leadership dimension, the evidence for which comes almost entirely through legislative speech, earns the higher dimension weight, see Figure 7. The left panel of the figure compares estimates for the voting dimension, which is the second dimension estimated with the heavily censored data (plotted along the vertical y -axis) while it corresponds with the first dimension of the uncensored estimates (graphed relative to the horizontal x -axis). Observations are labeled by party, and leaders' locations are in bold and circled. As one would expect, with less than $1/25^{th}$ of the voting data, recovery of the first dimension is far from perfect, but remarkably the imputed scores correlate highly, at more than 0.85, with the estimates

based on the full data set. The right hand panel compares estimates for the “leadership” dimension, which coincides with the first dimension based on the censored data, but with the second dimension based on the uncensored data set. In contrast with the voting dimension, the censored estimates correspond closely with their uncensored counterparts. Of course, the “leadership” dimension is driven mostly by words, and we did not censor those.

While this last exercise may seem a stunt, we note that in heavily whipped parliaments most legislators vote their parties, rather than their preferences (e.g., Kellerman, 2012), yet they still give speeches. In such settings we might use SFA to “bridge” between speeches actually given by members of a parliament to the votes that they would have cast had they not been “whipped”, anchoring the exercise by treating the votes of party principals as a genuine reflection of the leaders’ preferences.

Finally, we use words to bridge legislative and non-legislative actors. Whereas the previous analysis confronted an artificial inferential hurdle, in which we intentionally censored our voting data and compared speech among legislators, we are sometimes interested in political speakers who do not sit in the legislature at all. What can we learn about their preferences on the voting dimension by comparing their speech to that of Senators? We broach this question by incorporating word count data from unsigned editorials published during the two years that the 112th Congress was in session in the *New York Times*, the *Wall Street Journal*, and the *Washington Post*, using the same terms we employed in our analysis of the Senate. As the term data come from different venues, the Senate floor versus the penultimate page of the first section of a newspaper, the exercise is one of “out of sample prediction”. This leaves us with the question of whether the political meanings of the terms of discourse are the same in both venues. As a first approach to this issue, we treat the ideal points for both groups as coming from a mean-zero distribution. Results appear in Figure 8. We orient the dimension so that the Republicans have a positive value. The densities for the Republican and Democratic Senators are in the background, and the voting dimension legislator preferred outcomes are plotted as hatch marks along the x -axis. The results are largely as expected. If we treat the three sets of editorial boards as legislators who do not vote, we find the Wall Street Journal (**WSJ**) to the right of the *Washington Post* (**Wash Post**) and the *New York Times* (**NYT**) to its left. The distance between the *Wall Street Journal* and the *Washington Post* is about half the estimated distance between the *New York Times* and the *Post*.

Newspapers Placed on the Same Scale as Legislators

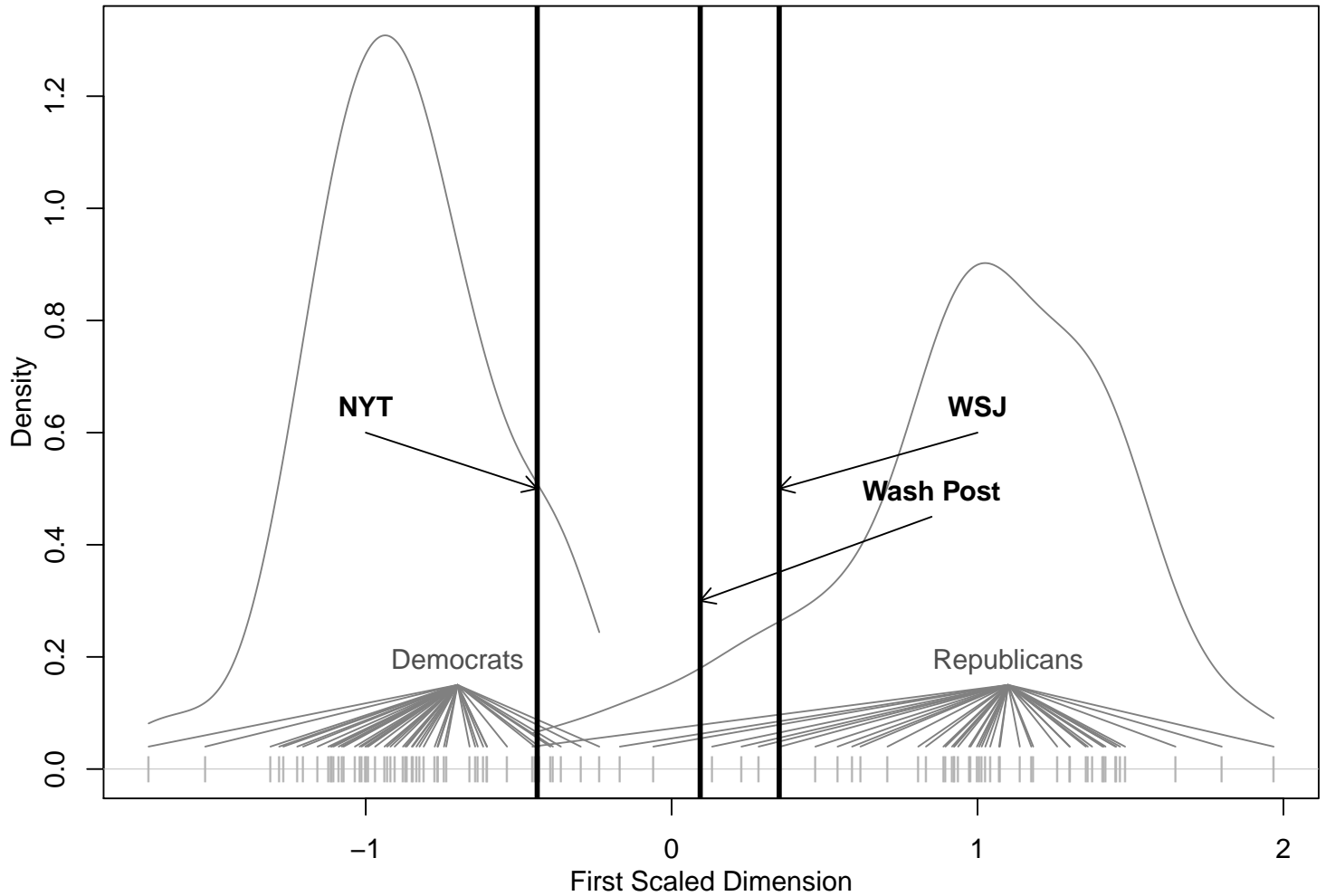


Figure 8: **Scaling newspaper editorials given only their text.** This figure presents the relative locations and differences between the ideal points for legislators and newspapers.

5 Conclusion

This paper offers insights both substantive and methodological. Substantively, we analyze legislative speech and roll call voting from eight recent sessions of the US Senate to reveal a consistent picture of a two dimensional Senate, with the first dimension coinciding with the voting dimension, while the second distinguishes leaders of both parties from the rank and file. Essentially the first dimension reflects the relative positions of Senators and the ideological affect of words, while the second dimension identifies party leaders.

If we consider only votes, our estimates echo the near consensus in the Congressional literature that recent voting behavior has been unidimensional. If in contrast we omit votes and model only speech, we find a larger number substantive dimensions. Our findings resonate with the longstanding assertion that the organization of Congress channels multidimensional policy preferences into a single voting dimension, *e.g.* Shepsle (1979) and Shepsle and Weingast (1981).

In order to generate our estimates, we develop a method, Sparse Factor Analysis (SFA) that encompasses both word choice and vote choice in terms of the same ideal points. We furthermore develop a statistical framework that allows us to estimate both individuals' most preferred outcomes and the underlying dimensionality of the joint word-vote space. The resulting methodology provides a close linkage between the choice-theoretic models of vote and word choice. This tight connection allows the extension of SFA to more complex decision scenarios (Clinton and Meirowitz, 2003; Signorino, 1999). Second, SFA allows the analyst to estimate the underlying number of latent dimensions, rather than having to impose dimensionality *a priori*.

While SFA is designed to analyze individuals who both speak and cast votes, it allows us to impute policy preferences to non-voting political speakers, a potential we illustrated for the case of newspaper editorial boards. This may prove useful in confronting the perennial research problem of imputing the preferred policy outcomes of legislative candidates. While analysts can impute the ideology of victorious candidates from their subsequent congressional conduct, as they can infer the leanings of defeated incumbents from their previous voting records, measuring the preferences of defeated challengers has proven to be a more elusive goal. Yet every challenger spends time and energy generating political speech. SFA offers the possibility of imputing the most preferred policy such a candidate would have pursued had he been elected.

We hope the approach in this paper also finds purchase beyond the US Congress. For example, in strong party systems where votes are relatively uninformative, words may be used to help clarify the within-party variance in ideal points. We are currently exploring applications of the method in situations where voting is not perfectly reflective of underlying individual preference. We are also working to extend SFA to a dynamic setting, allowing an over-time scaling of the same actors (Martin and Quinn, 2002). The SFA framework allows extension to a broad array of substantive questions, and we invite legislative

scholars to join us in pursuing its potential.

References

- Aldrich, John. 1995. *Why Parties?: The Origin and Transformation of Political Parties in America*. American Politics and Political Economy Series University of Chicago Press.
- Aldrich, John, Jacob Montgomery and David Sparks. 2014. "Polarization and Ideology: Partisan Sources of Low Dimensionality in Scaled Roll Call Analyses." *Political Analysis* .
- Austen-Smith, David. 1990. "Information Transmission in Debate." *American Journal of Political Science* 34(1):124–152.
- Blaydes, Lisa, Justin Grimmer and Alison McQueen. 2014. "Mirrors for Princes and Sultans: Advice on the Art of Governance in the Medieval Christian and Islamic Worlds." Working Paper.
- Blei, David M., Andrew Y. Ng and Michael I. Jordan. 2003. "Latent dirichlet allocation." *J. Mach. Learn. Res.* 3:993–1022.
- Bonica, Adam. 2014. "Mapping the Ideological Marketplace." *American Journal of Political Science* 58(2):367–386.
- Cattell, Raymond Bernard. 1978. *The Scientific use of Factor Analysis in Behavioral and Life Sciences*. New York: Plenum.
- Clinton, Joshua and Adam Meirowitz. 2003. "Integrating voting theory and roll call analysis: A framework." *Political Analysis* 11:381–396.
- Clinton, Joshua, Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98:355–370.
- Collins, Michael, Sanjoy Dasgupta and Robert E. Schapire. 2002. "A Generalization of Principal Component Analysis to the Exponential Family." In NIPS.
- Cox, Gary and Matthew McCubbins. 1993. *Legislative leviathan: Party government in the House*. University of California Press.

- Diaconis, Persi, Sharad Goel and Susan Holmes. 2008. "Horseshoes in Multidimensional Scaling and Local Kernel Methods." *Annals of Applied Statistics* 2(3):777–807.
- Donoho, David L. and Iain M. Johnstone. 1994. "Ideal Spatial Adaptation by Wavelet Shrinkage." *Biometrika* 81(3):425–455.
- Epstein, Lee, Jeffrey Segal and Harold Spaeth. 2001. "The Norm of Consensus on the US Supreme Court." *American Journal of Political Science* 45(2):362–377.
- Fenno, Jr., Richard F. 1977. *Congress, Keystone of the Washington Establishment*. New Haven: Yale University Press.
- Fenno, Jr., Richard F. 1978. *Home Style: House Members in their Districts*. Boston: Little, Brown.
- Gentzkow, Matthew and Jesse M. Shapiro. 2010. "What Drives Media Slant? Evidence From U.S. Daily Newspapers." *Econometrica* 78(1):35–71.
- Gerrish, Sean and David M. Blei. 2012. How They Vote: Issue-Adjusted Models of Legislative Behavior. In *Advances in Neural Information Processing Systems 25*, ed. F. Pereira, C.J.C. Burges, L. Bottou and K.Q. Weinberger. Curran Associates, Inc. pp. 2753–2761.
- Gilligan, Thomas and Keith Krehbiel. 1987. "Collective decisionmaking and standing committees: An informational rationale for restrictive amendment procedures." *Journal of Law, Economics, & Organization* 3(2):287–335.
- Golub, Gene and Charles van Loan. 1996. *Matrix Computations*. Johns Hopkins Studies in Mathematical Sciences 3rd ed. Baltimore: Johns Hopkins University Press.
- Grimmer, Justin. 2010. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis* 18(1):1–35.
- Grimmer, Justin and Brandon Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21(3):267–297.
- Hahn, P. Richard, Carlos M. Carvalho and James G. Scott. 2012. "A sparse factor analytic probit model for congressional voting patterns." *Journal of the Royal Statistical Society, Series A* 61(4):619–635.

- Heckman, James J. and James M. Snyder Jr. 1997. "Linear Probability Models of the Demand for Attributes with an Empirical Application to Estimating the Preferences of Legislators." *The RAND Journal of Economics* 28:S142–S189.
- Hill, Kim Quaile and Patricia A. Hurley. 2002. "Symbolic Speeches in the U.S. Senate and Their Representational Implications." *The Journal of Politics* 64(1):219–231.
- Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. Wiley.
- Jackman, Simon and Shawn Trier. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 52(1):201–17.
- Kellerman, Michael. 2012. "Estimating Ideal Points in the British House of Commons Using Early Day Motions." *American Journal of Political Science* 56(3):757–771.
- Krehbiel, Keith. 1998. *Pivotal Politics*. University of Chicago Press.
- Ladha, Krishna. 1991. "A Spatial Model of Legislative Voting with Perceptual Error." *Public Choice* 68:151–74.
- Lauderdale, Benjamin and Alexander Herzog. 2014. "Measuring Political Positions from Legislative Debate Texts on Heterogenous Topics." Working Paper.
- Lauderdale, Benjamin and Tom Clark. 2014. "Scaling Politically Meaningful Dimensions Using Texts and Votes." *American Journal of Political Science* Forthcoming.
- Maltzman, Forrest and Lee Sigelman. 1996. "The Politics of Talk: Unconstrained Floor Time in the U.S. House of Representatives." *The Journal of Politics* 58(3):819–30.
- Martin, Andrew and Kevin Quinn. 2002. "Dynamic Ideal Point Estimation via Markov chain Monte Carlo for the U.S. Supreme Court, 1953–1999." *Political Analysis* 10(2):134–153.
- Mayhew, David. 1974. *Congress: The Electoral Connection*. Yale: New Haven Press.
- McKelvey, Richard. 1976. "Intransitivities in Multidimensional Voting Models and Some Implications for Agenda Control." *Journal of Economic Theory* 12(475–482).

- McKelvey, Richard and W. Zavoina. 1975. "A statistical model for the analysis of ordered level dependent variables." *Journal of Mathematical Sociology* 4(103–20).
- Mohamed, Shakir, Katherine Heller and Zoubin Ghahramani. 2008. "Bayesian Exponential Family PCA." In NIPS.
- Monroe, Burt, Michael Colaresi and Kevin Quinn. 2008. "Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict." *Political Analysis* 16(4):372–403.
- Murray, Jared S., David B. Dunson, Lawrence Carin and Joseph E. Lucas. 2013. "Bayesian Gaussian Copula Factor Models for Mixed Data." *Journal of the American Statistical Association* 108(502).
- Neal, Radford. 2011. MCMC Using Hamiltonian Dynamics. In *Handbook of Markov Chain Monte Carlo*, ed. Steve Brooks, Andrew Gelman, Galin Jones and Xiao-Li Meng. CRC Handbooks of Modern Statistical Method Chapman and Hall.
- Park, Trevor and George Casella. 2008. "The bayesian lasso." *Journal of the American Statistical Association* 103(482):681–686.
- Poole, Keith and Howard Rosenthal. 1997. *Congress: A Political Economic History of Roll Call Voting*. New York: Oxford University Press.
- Quinn, Kevin M. 2004. "Bayesian Factor Analysis for Mixed Ordinal and Continuous Responses." *Political Analysis* 12(4):338–353.
- Quinn, Kevin M., Burt L. Monroe, Michael Colaresi, Michael H. Crespin and Dragomir R. Radev. 2010. "How to Analyze Political Attention with Minimal Assumptions and Costs." *American Journal of Political Science* 54(1):209–228.
- Roberts, Molly, Brandon Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Gadarian, Bethany Albertson and David Rand. 2014. "Structural topic models for open ended survey responses." *American Journal of Political Science* 58:1064–1082.
- Schofield, Normal. 1977. "Transitivity of Preferences on a Smooth Manifold." *Journal of Economic Theory* 14(149–72).

Shepsle, Kenneth. 1979. "Institutional Arrangements and Equilibrium in Multidimensional Voting Models," *American Journal of Political Science* 23:23–57.

Shepsle, Kenneth and Barry Weingast. 1981. "Structure-Induced Equilibrium and Legislative Choice." *Public Choice* 37:503–519.

Signorino, Curtis. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *American Political Science Review* 93(2):279–297.

Sim, Yanchuan, Bryan Routledge and Noah Smith. 2014. "The Utility of Text: The Case of Amicus Briefs and the Supreme Court."

Slapin, Jonathan B. and Sven-Oliver Proksch. 2008. "A Scaling Model for Estimating Time Series Party Positions from Texts." *American Journal of Political Science* 52(3):705–722.

Wang, Eric, Esther Salazar, David Dunson and Lawrence Carin. 2013. "Spatio-Temporal Modeling of Legislation and Votes." *Bayesian Analysis* 8(1):233–268.

Abstract: 147 Words

Body: 10,063 Words