Measuring Social Ties from Roll Call Votes: A Fused Latent Factor and Social Network Approach

Hane Lee* Andrew Davison Zhiliang Ying

Department of Statistics, Columbia University

Draft: November 20, 2024

Abstract

Congressional literature suggests that the motivations behind roll call votes are complex, spanning the legislator's ideology, party strategies, and social influences. In terms of methodology, latent factor models have dominated roll call analysis, where the estimated "ideal points" are interpreted as the legislators' partisan-ideological positions, but these models do not account for partisan or social motivations behind the votes. On the other hand, some researchers have explored the social influence behind these votes using network models, but this approach often overlooks the role of ideology or parties. We address this gap by integrating the partisan-ideological and social approaches through a fused latent factor and social network model. This model decomposes the effects of partisan-ideology and social connections on roll call votes while giving priority to the former. Additionally, our method provides a direct measurement of social ties from roll call votes, rather than relying on proxies such as cosponsorship to first estimate the social effect and later make connections to political outcomes. We apply our model to the 101st Senate and find that the model successfully decomposes ideology and partisanship from social ties. The estimated social network captures notable friendships and geographical communities. We also demonstrate that cosponsorship and shared committee membership, commonly viewed as indicators of social connections, are either closely aligned with the legislator's revealed partisan-ideological preferences or have minimal legislative impact.

Keywords: Roll Call Votes, Social Network, Latent Factor Models, Item Response Theory, Fused Models

*hane.lee@columbia.edu

1 Introduction

Roll call votes have long assisted congressional research as some of the most visible and consequential outputs of individual legislators, and understanding the attitude underlying these votes has been of crucial academic interest. Many researchers of electoral representation attributed the votes to the personal preferences of the legislator (Miller and Stokes 1963; Mayhew 1974), reflecting both their individual attitude and perception of public preferences. Later, researchers of party politics point out that the revival of parties and bipartisan leadership in the 1980s led to the increase of party voting (Rohde 1991; Snyder and Groseclose 2000; Lee 2009).

Methodological approaches for roll call analysis have favored the ideological hypothesis, especially since the introduction of NOMINATE (Poole and Rosenthal 1985) and following latent factor models (Heckman and Snyder 1997; Clinton, Jackman, and Rivers 2004). Under the assumption that legislators are earnest ideological actors, the estimated latent factor, or "ideal point", of each legislator is interpreted as their true ideological position on a Downsian spatial model. This methodology has facilitated an ideological interpretation of the Congress; for instance, McCarty, Poole, and Rosenthal (2016) demonstrated the increasing divergence between the ideal points of Democratic and Republican legislators since the mid-1970s, suggesting that this gap evidences ideological polarization in the Congress.

Unfortunately, latent factor models do not provide an entirely satisfactory explanation of roll-call votes. For instance, DW-NOMINATE, a dynamic and weighted extension of NOMINATE, accurately predicts only 83% of the votes with one latent factor, interpreted as "left-right economic views"; a second factor, often described as "cross-cutting salient issues", narrowly raises accuracy by 2% (Everson et al. 2016). Poole and Rosenthal (2000) acknowledge that latent factor models alone do not give a complete explanation of roll-call votes because many external factors other than latent political or economic views, such as partisanship, strategic voting, and social ties, also influence the votes.

Researchers have long acknowledged that social ties heavily influence collective decision making processes among legislators, both on the federal (Baker 1980; Caldeira and Patterson 1987) and local levels (Arnold, Deen, and Patterson 2000). Network analysis was used to measure these social ties and their influence on roll call votes. Fowler (2006) argued that cosponsorship often reflects social proximity between legislators, and better connected legislators on the cosponsorship network tend to gain more votes for the bills that they sponsor. Kim, Barnett, and Kwon (2010) examined 6 different social networks, each constructed from cosponsorship, party membership, shared PAC donation, geographical contiguity, shared committee membership, and internet hyperlinks, and find that the first four predict roll call votes.

While network analysis has offered evidence on the existence of social effects in the Congress, there are many limitations to existing applications. First, these networks are constructed from proxies that may result in social ties rather than a direct measurement of social ties, so the existence of an edge between two legislators does not guarantee a close social relationship between them (Kirkland and Gross 2014). Second, most of these networks are constructed from only one source (e.g. cosponsorship, committee membership) and fail to capture the heterogeneity of social ties in the congress. Each network may also overlook social ties that originate from other sources; for instance, the network constructed from shared committee membership will not capture social ties formed from geographical contiguity. Third, when these social networks are connected to political outcomes, these outcomes are attributed entirely to social connections, though ideology and partisanship are often more decisive factors.

To address these limitations, we aim to directly measure social connections from roll call votes while controlling for of partisanship and ideology. By employing a fused latent factor and network model that integrates the ideological and social approaches to roll call analysis, we simultaneously estimate ideal points and a social network from roll call votes. This model decomposes the effects of partisanship and ideology from those of social ties on roll call votes.

This paper has three main contributions. First, to the best of our knowledge, our paper is the first to provide a direct measure of social ties from roll call votes. Most existing applications construct potential social networks from proxy connections and then measure its influence on political outcomes. Second, we integrate two models of roll call analyses, the ideological latent factor model and the social network model, on both theoretical and methodological levels. Assuming that social ties are secondary influences to ideology and partisanship, the fused model that we suggest can decompose these effects. Third, we demonstrate that geographical proximity plays a prominent role, while cosponsorship and shared committee membership, traditionally regarded as influential social connections, have limited impact on roll call behavior outside of the partisan-ideological alignment.

The rest of the paper is organized as follows. In section 2, we review and integrate existing theoretical models of roll call behavior that focus on ideology, partisanship, and social ties. Next, in section 3, we introduce a fused latent and graphical model that reflects the combined model. In section 4, we apply this model to the 101st Senate. Section 5 concludes.

2 Integrating the Partisan-Ideological and Social Theories of Roll Call Behavior

2.1 Models of Ideology, Party, and Roll Call Behavior

Electoral studies on representation have extensively discussed theories of roll call behavior, where roll call votes are often modeled as a function of a legislator's own attitude and their constituency's preferences. Miller and Stokes (1963) suggested that constituents control their Representative's roll call behavior through two pathways: electing a legislator whose attitudes that are most aligned with their own, and providing a perception of preferences that Representatives must respond to in order to secure re-election. Mayhew (1974) argued that legislators, as "single-minded re-election seekers", use roll call votes to signal their alignment with popular opinion. Arnold (1990) noted that this signaling is stronger for salient issues, as legislators aim to respond to the potential interests of an inattentive public.

Roll call behavior is also discussed by the literature on party politics. Rohde (1991) demonstrated that while the 1970s party reform decentralized the House, the post-reform electoral realignment increased party homogeneity, leading to increased party voting in the late 70s and 80s. Snyder and Groseclose (2000) examined the extent of party influence on roll call behavior through regression, while controlling for individual preferences estimated from lopsided roll calls, which are less influenced by parties. They found strong party influence on roll call votes between 1871 and 1998. Lee (2009) discovered that non-ideological issues, comprising more than half the Senate bills from 1981 to 2004, showed partisan divide, highlighting the substantial influence of parties alongside personal ideological preferences.

A central debate in this literature concerns whether individual preferences or parties exert more influence over the legislative decisions of a legislator. Krehbiel (1993) examined the standing committee formation and conferee appointments in the 99th Congress and pointed to the lack of "significant party behavior" that is consistent with party objectives and is independent of, or even in conflict with, personal preferences. He argues that, therefore, parties are not central to legislative decisions. In response, Aldrich and Rohde (2000) contended that not all issues are relevant to partisan interests. They offered the theory of "conditional party government"—empirically supported by Rohde (1991) through roll call votes in the House—which suggests that parties pursue party interests only when members are unified.

While the literature has often divided focus between individual and constituency preferences or party affiliation as the primary drivers for roll call decisions, these factors cannot be easily separated. Many researchers acknowledge the high correlation between individual preferences, constituency preferences, and party affiliation (Snyder and Groseclose 2000), which result in complex, overlapping effects on roll call behavior. Arnold (1990) argued that legislators are most influenced by coalition leaders—such as committee leaders, party leaders, congressional staff members, the president, and interest group leaders—as well constituency and personal preferences. Ansolabehere, Snyder, and Stewart (2001) compared surveys of House members during the 1996 and 1998 elections with their roll call votes, finding that both individual preferences and party are significant predictors of roll call behavior, with party effects independent of individual preferences appearing in approximately 40% of roll calls. Given the combined theories and empirical evidence, we adopt the following assumption for our analysis:

A1. *Individual preferences and party policies are the primary motivations behind roll call votes.*

The heterogeneity of influences on roll call votes, combined with the high correlation between the sources, raises important questions regarding interpretation of latent factor models in roll call analysis. Latent factor models detect patterns in data, where each orthogonal factor represents an underlying construct that shapes the observed responses. In the case of roll call votes, the estimated ideal points of legislators are often equated with their ideological positions (e.g. Canes-Wrone, Brady, and Cogan 2002). This connection originates from the Downsian spatial model (Downs 1957) where voters, as rational agents, are positioned at their ideal point on an ideological space and choose the party closest to their ideal point. This framework was further methodologically formalized by NOMINATE (Poole and Rosenthal (1985)) and subsequent latent factor models that were developed to estimate the ideal points, and ultimately the ideologies, of individual legislators.

Though the dominant narrative often translates the estimated latent factors as ideological positions, this interpretation is neither inherent nor predetermined. Once the factors are estimated, they must be manually observed and labeled by the researcher, and such interpretations cannot be easily validated. Several researchers have highlighted potential issues with this approach. Caughey and Schickler (2016) pointed out that the two-dimensional latent factors of NOMINATE do not maintain consistent ideological meaning over time. Additionally, though the first factor of NOMINATE is often interpreted as individual economic ideology, the clear separation of ideal points across party lines suggests a strong alignment between individual "ideology" and party policies. Recognizing this alignment and the prevalence of strategic party voting, many researchers instead attempt to estimate the legislators' true individual preferences independent of the party by using a selection of votes where party influence is minimal (Snyder and Groseclose 2000; Minozzi and Volden 2013).

More recently, Ramey, Klingler, and Hollibaugh (2017) observed that many scholars interpret NOMINATE-type measures as legislators' "revealed preferences", which reflects not only individual preferences but also party influences. Based on **A1** and this interpretation, we expect that the latent factor models will capture a combination of individual preferences and partisanship as dominant factors, and we will later verify if this is the case.

2.2 Modeling Social Ties as Secondary Influence

Formal theory linking congressional social connections to legislative impact was first developed by Kirkland (2011). Drawing from social network theory, he proposed a model of influence diffusion, suggesting that legislators, as goal-oriented agents, utilize social interactions to gain informational influence and ultimately reach legislative success.

However, research on social ties in Congress predates this formal theory. Caldeira and Patterson (1987) argued that social connections provide the foundation for partisan dynamics, bargaining, vote cue communication, and other political activities. In a regression analysis of roll call data from the 120th Ohio House of Representatives (1993-1994), Arnold, Deen, and Patterson (2000) found that while partisanship was the strongest predictor of shared roll call behavior, friendship was a more significant predictor than ideology, gender, or geographical proximity. Fowler (2006) explored the social implications of the cosponsorship network and demonstrated that connectedness within this network is positively correlated with bill passage, even when controlling for individual ideology and partisanship. More recently, Fong (2020) found that legislators often rely on cues from trusted peers when dealing with bills outside their area of expertise. Similarly, Battaglini, Sciabolazza, and Patacchini (2020, 2023) demonstrated that social ties influence legislative effectiveness and voting behavior, particularly abstentions.

Research on social connections in the Congress predates this formal theory. Caldeira and Patterson (1987) suggested that social ties supply the foundation for partisan dynamics, bargaining, communication of vote cues, and other political activities. In a regression analysis of roll call data from the 120th Ohio House of Representatives (1993-4), Arnold, Deen, and Patterson (2000) observed that while partisanship was the strongest predictor of shared roll call behavior, friendship was a stronger predictor than ideology, gender, or geographical proximity. Fowler (2006) proposed a social interpretation of the cosponsorship network and demonstrated that connectedness on this network is positively correlated bill passage, even when controlling for individual ideology and partisanship. More recently, Fong (2020) showed that legislators often rely on cues from trusted peers when voting for bills outside their area of expertise. Battaglini, Sciabolazza, and Patacchini (2020, 2023) showed that social connections influence legislative effectiveness and abstentions.

Researchers agree that social ties significantly influence roll call behavior, but the extent of this influence remains debated, especially in relation to legislators' individual preferences and partisanship. As discussed in the previous section, the importance of these two factors in shaping roll call behavior is wellestablished, and scholars examining social connections agree that individual preferences and partisanship are central, and perhaps more dominant, factors in roll call decisions, which may even drive the formation of social connections (Kirkland 2011; Bratton and Rouse 2011; Fowler 2006; Burkett 1997). Drawing from these insights, we make the following assumption.

A2. *The influence of social connections on roll call behavior is secondary to that of the legislators' individual preferences and party strategies.*

However, there are several other important factors that shape social connections, such as geography, gender, and religion, which do not always align with ideological or partisan divides. Moreover, the significance social ties becomes most evident when they operate independently of, or even challenge, the more dominant influences of ideology and partisanship. With this in mind, we introduce the following assumption.

A3. *Social influences are most apparent in votes that diverge from a legislator's typical partisan-ideological voting patterns.*

3 A Fused Latent Factor and Social Network Model

In this section, we aim to empirically integrate the partisan-ideological and social approaches to roll call analysis. We assumed that partisan-ideological patterns of voting are dominant, social influences are secondary, and legislators cast votes atypical of their central patterns due to social influences. To operationalize these theoretical assumptions, we adopt a fused latent factor and social network model developed by Chen et al. (2016). In this model, votes are primarily attributed to latent factors, while any remaining variation is cast to a social network. The latent factors capture the core partisan-ideological voting behavior, while the network explains deviations from this pattern though correlations between legislators' atypical votes. We will later interpret these correlations as social connections.

3.1 The M2PL-Ising Model

The multidimensional two-parameter logistic (M2PL)-Ising model, developed by Chen et al. (2016), brings together two models with historically different usage. The first is a common item response theory model for measuring latent traits given binary responses. Given a two-year term of congress with J legislators and N bills, with the jth legislator's vote for the nth bill denoted by $Y_{jn} \in \{0, 1\}$, the M2PL model defines the probability of a "Yea" vote ($Y_{jn} = 1$) as a logistic distribution:

$$P(Y_{jn} = y_{jn} | \boldsymbol{\theta}_n, \mathbf{a}_j, d_j) = \frac{\exp\Big\{(\mathbf{a}_j^\top \boldsymbol{\theta}_n + d_j) y_{jn}\Big\}}{1 + \exp\Big\{(\mathbf{a}_j^\top \boldsymbol{\theta}_n + d_j) y_{jn}\Big\}} \propto \exp\Big\{(\mathbf{a}_j^\top \boldsymbol{\theta}_n + d_j) y_{jn}\Big\}. \tag{1}$$

Here, $\boldsymbol{\theta}_n \in \mathbb{R}^K$ denotes the $K > 1$ dimensional latent factors of the bill n, $\mathbf{a}_j \in \mathbb{R}^K$ the factor loadings or the discrimination parameter of the jth legislator, and d_j the difficulty parameter of the jth legislator. Proceeding with equation (1), a local independence assumption is typically made to obtain the joint conditional distribution, with $A_{J\times K} = (\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_J)^\top$ and $\mathbf{d} = (d_1, d_2, \cdots, d_J)^\top$:

$$P(\mathbf{Y}_n = \mathbf{y}_n | \boldsymbol{\theta}_n) \propto \exp\Big\{\boldsymbol{\theta}_n^\top A^\top \mathbf{y}_n + \mathbf{d}^\top \mathbf{y}_n\Big\}.$$

The M2PL model with $K = 2$ is analogous to the two-dimensional latent factor model of NOMINATE, with the key distinction being that NOMINATE uses a probit link to model $P(Y_{jn} = y_{jn})$, whereas M2PL uses a logit link (Clinton, Jackman, and Rivers 2004).

The second component is the Ising model, an exponential family graphical model originating from physics, which has been increasingly applied to social network settings to capture the conditional graphical dependence structure among a set of random variables (Stauffer 2008). Given binary random variables, each variable is associated with a vertex, and the undirected graphical structure among variables is mapped on a symmetric matrix $S = (s_{ij})_{J\times J}$ where s_{ij}—which can take positive, negative, and 0 values—indicates the "strength" of the edge between two variables. In the context of roll call votes, a larger s_{ij} indicates a higher probability that the ith and jth legislators will simultaneously vote "Yea". The Ising model is defined as

$$P(\mathbf{Y}_n = \mathbf{y}_n) = \frac{1}{z(S)} \exp\left\{\frac{1}{2}\mathbf{y}_n^\top S \mathbf{y}_n\right\},$$

where $z(S)$ is the normalizing constant.

Taking $L = AA^\top$ with factor loading matrix A of the M2PL model and graph S of the Ising model, the M2PL-Ising model is a fusion of the two models, with the joint distribution given by

$$p(\mathbf{y}_n | L, S) = \frac{1}{z(L, S)} \exp\left\{\frac{1}{2}\mathbf{y}_n^\top (L + S)\mathbf{y}_n\right\}, \tag{2}$$

where $z(L, S)$ is again the normalizing constant. L captures the dependencies among the votes that are driven by the latent structure, which we interpret as ideology and partisanship, while S represents the network dependencies, which we interpret as social effects.

3.2 Estimation

The normalizing constant $z(L, S)$ from the model in equation (2) is computationally intractable. Therefore, to estimate L and S, Chen et al. (2016) propose a pseudolikelihood estimator. For jth legislator and nth bill, the pseudolikelihood is defined as

$$\begin{aligned}
\mathcal{L}_{jn}(L, S; \mathbf{y}_n) &\equiv P(Y_{jn} = 1 | \mathbf{Y}_{-j,n} = \mathbf{y}_{-j,n}, L, S) \\
&= \frac{\exp\left\{\frac{l_{jj}+s_{jj}}{2} + \sum_{i \neq j}(l_{ij} + s_{ij}) y_{in}\right\}}{1 + \exp\left\{\frac{l_{jj}+s_{jj}}{2} + \sum_{i \neq j}(l_{ij} + s_{ij}) y_{in}\right\}}, \\
\mathcal{L}(L, S) &= \prod_{n=1}^{N} \prod_{j=1}^{J} \mathcal{L}_{jn}(L, S; \mathbf{y}_n).
\end{aligned} \tag{3}$$

The pseudolikelihood in equation (3) depends on parameters L and S only through $L+S$. Therefore, additional assumptions are necessary to identify L and S from the data. To address this, Chen et al. assume that most of the dependence in Y is captured by the latent factors, and the remaining dependence is attributed to the graphical structure. This leads to the specification of L as a low-rank matrix and

S as a symmetric sparse matrix. This specification has important implications for roll call analysis as it allows a decomposition of latent and network effects with the latent effect taking precedence, consistent with our theoretical assumptions A1-3.

We estimate L and S by minimizing the pseudolikelihood with two regularization terms: the nuclear norm of L to induce the low rank of L, and an L_1 regularization term of S to induce element-wise sparsity (Wright and Ma 2022), with ρ and γ as tuning parameters. We additionally assume that L is positive semi-definite and S is symmetric.

$$(\hat{L}, \hat{S}) = \arg\min_{L,S} \Big\{ -\frac{1}{N} \log\{\mathcal{L}(L, S)\} + \rho \|L\|_* + \gamma \sum_{i \neq j} |s_{ij}| \Big\}$$

To fit the pseudolikelihood estimator, we follow Chen et al. and implement the alternating direction method of multipliers (ADMM) algorithm (Boyd et al. 2011), a popular method for convex optimization. For the objective above, this takes following form:

$$\min_{\mathbf{x},\mathbf{z}} f(\mathbf{x}) + g(\mathbf{z}) \quad \text{s.t. } \mathbf{x} = \mathbf{z},$$

$$\begin{aligned}
\text{where } \mathbf{x} &= (M, L, S), \\
\mathbf{z} &= (\tilde{M}, \tilde{L}, \tilde{S}), \\
f(\mathbf{x}) &= \begin{cases} h(M) + \rho\|L\|_* + \gamma \sum_{i!=j} |s_{ij}|, & \text{if } L \succeq, S^\top = S \\ \infty, & \text{otherwise} \end{cases} \\
g(\mathbf{z}) &= \begin{cases} 0, & \text{if } M^\top = M, M = L + S \\ \infty, & \text{otherwise} \end{cases}
\end{aligned}$$

We execute the following updates until all primal and dual residuals meet the convergence criteria, with step size λ:

$$\mathbf{x}^{k+1} = \arg\min_{\mathbf{x}} \left\{ f(\mathbf{x}) + \frac{1}{2\lambda} \left\| \mathbf{x} - (\mathbf{z}^k - \mathbf{u}^k) \right\|_2^2 \right\},$$

$$\mathbf{z}^{k+1} = \arg\min_{\mathbf{z}} \left\{ g(\mathbf{z}) + \frac{1}{2\lambda} \left\| \mathbf{z} - (\mathbf{z}^k - \mathbf{u}^k) \right\|_2^2 \right\},$$

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \mathbf{x}^{k+1} - \mathbf{z}^{k+1}.$$

This algorithm was implemented in Python, with the code available on GitHub. We fit multiple models $\mathcal{M}^{\rho,\gamma}$ over ranges of the tuning parameters ρ, γ and select the model that minimizes the Bayesian Information Criterion (BIC),

$$\mathrm{BIC}(\mathcal{M}^{\rho,\gamma}) = -2 \log \mathcal{L}(\hat{L}^{\rho,\gamma}, \hat{S}^{\rho,\gamma}) + |\mathcal{M}^{\rho,\gamma}| \log N.$$

4 Application to the 101st Senate

Social connections play a more central role in the legislative process of the Senate than the House (Sinclair 2016, Chapter 3), partly due to the Senate's smaller size and less hierarchical structure. More importantly, unique structural features of the Senate such as unanimous consent agreements and filibusters require supermajority cooperation that necessitates "relation-based legislating" that relies on informal negotiations beyond formal rules and procedures (Wawro and Schickler 2007).

We apply the M2PL-Ising model to roll call votes from the 101st Senate (1989-90) with $N = 638$ bills and $J = 101$ senators (with one change in membership during the session, for Hawaii), where Democrats held majority with 55 seats. We first examine the estimated latent factors to confirm whether the model captures the partisan-ideological voting patterns. Next, we will examine the estimated social network to identify patterns of social connections.

4.1 Estimated Latent Factors

In section 2, we assumed that partisan and individual ideological preferences drive the dominant patterns of voting. Here, we verify whether our fitted latent factor estimations captures these patterns.

We restricted our model selection to those with latent dimensions of 2 or higher for consistency with

	p_{D}	p_{R}
$\hat{\boldsymbol{\theta}}_1$	0.985	0.038
$\hat{\boldsymbol{\theta}}_2$	0.153	0.990

Table 1: Correlation between the estimated latent factors $(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2)$ and the proportion of Democrat/Republican yea votes per bill $(p_{\mathrm{D}}, p_{\mathrm{R}})$.

	$\hat{A}_1$	$\hat{A}_2$
$\boldsymbol{\theta}_{\mathrm{NOM1}}$	-0.93209468	0.70891187

Table 2: Correlation between the first dimension of NOMINATE scores (θ_{NOM1}) and the two dimensions of the esimated discrimination parameters ($\hat{A}_1$ and $\hat{A}_2$).

other latent factor models that typically use a two-dimensional factor space. Our final choice is the model with the lowest BIC and $\mathrm{rank}(\hat{L}) = 2$. Following Chen et al., we extract the factor loadings $\hat{A}$ from $\hat{L} = \hat{A}\hat{A}^\top$ using varimax rotation and also estimate $\hat{\boldsymbol{\theta}} = \hat{A}^\top Y$, the latent factor of bills.

We find in Table 1 that the first dimension of the estimated latent factors $(\hat{\boldsymbol{\theta}}_1)$ is highly correlated (0.985) with the proportions of "Yea" votes among Democrats for each bill (p_{D}). Similarly, the second dimension of factors $(\hat{\boldsymbol{\theta}}_2)$ is highly correlated (0.990) with the proportion of "Yea" votes among Republicans for each bill (p_{R}). These high correlations demonstrate that the estimated factors successfully capture patterns of partisan voting.

In our application, $\boldsymbol{\theta}$ indicate the latent factors of the bills, in contrast to existing latent factor applications where $\boldsymbol{\theta}$ typically corresponds to the ideal points of the legislators. While we do not directly estimate the ideal points of senators, we can use A, the discrimination parameters of senators, as a proxy to examine the ideal points of senators. Two features of A make it a useful proxy for senator latent factors: the sign of the slope, which indicates whether the senator prefers bills aligned with Democratic or Republican preferences, and the magnitude of the slope, which reflects the degree of partisan division of the senator's voting patterns.

The plot of $\hat{A}$ in Figure 1 indicate a significant degree of partisan separation. now we examine whether the model also reflects the ideological preferences of senators from previous applications. Table 2 presents

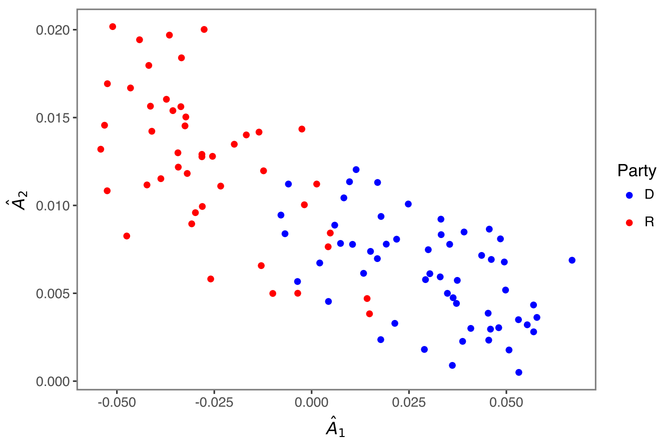

Figure 1: Plot of $\hat{A}$, where $\hat{A}_1$ refers to the first dimension of $\hat{A}$ and $\hat{A}_2$ the second. Blue dots represent Democrats and red dots represent Republicans.

the correlation between the two dimensions of $\hat{A}$ and the first dimension of NOMINATE scores, $\boldsymbol{\theta}_{\text{NOM1}}$. The absolute value of the correlations is high between both dimensions of $\hat{A}$ and $\boldsymbol{\theta}_{\text{NOM1}}$, indicating that the model captures a significant portion of the patterns in NOMINATE, up to a rotation. Therefore, we conclude that this model successfully captures the ideological patterns of voting that appear in existing latent factor applications.

Tables 1 and 2 demonstrate that the latent factor section of the fused model captures both partisanship and ideology, supporting the "revealed preferences" interpretation of latent factors in roll call analysis (Ramey, Klingler, and Hollibaugh 2017).

4.2 Estimated Social Network

The core partisan and ideological patterns of voting have been captured by L, and now we explore the remaining unexplained variation in votes through S. The fitted graphical matrix $\hat{S}$ for the selected model has sparsity 0.141 with 678 edges, 578 positive and 100 negative. We first explore notable connections in the network for preliminary analysis. Then, we examine whether $\hat{S}$ reflects connections from geographical proximity, cosponsorship, and shared committee membership.

4.2.1 Notable edges in $\hat{S}$

Partisan divide is evident in $\hat{S}$, where a majority of the positive edges (478/578) link senators from the same party, and a majority of the negative edges (75/100) link senators from different parties. furthermore, same-party edges are stronger on average (0.315) than different-party edges (0.214). A larger edge value $\hat{s}_{ij}$ indicates a stronger correlation between the atypical votes of senators i and j.

In the context of polarization, across-party positive edges signify favorable connections that formed despite the tension between parties, when, as former Minority Leader Bob Michel commented, "It's difficult at times to go out on the floor and just beat the hell out of your opposition and then expect them within a half hour to sit down and have a rational discussion" (Pianin 1991). We observe connections with the strongest edges in Table 3. Most of the pairs display geographical contiguity, with two pairs of senators from the same or nearby states. Many of these senator pairs also show an above-average (80.66) number of consponsored bills during the 2-year period. The most notable connections, however, are senators Joe Biden and Arlen Specter, who were famously good friends (Klein 2009), and senators Larry Pressler, James Exon, and Tom Harkin, who were veterans in addition to representing neighboring states.

Next, Table 4 lists the 10 strongest edges between senators from the same party. A prominent pattern across the pairs is the geographical proximity between the senators' constituencies. 6 pairs of senators represent the same state, while some others represent neighboring states, such as Iowa and South Dakota, or Texas and Oklahoma.

Table 5 lists senators who are most connected with the largest number of edges. The senators who are most connected within the party tend to have been in office for a long time or had leadership roles. Senator John Heinz, for instance, had recently stepped down as former chairman of the National Republican Senatorial Committee after two terms; Senator Jake Garn was first elected into office in 1974, and he retired in 1992 shortly after the 101st Congress. On the other hand, senators with the most out-of-party connections tend to be centrists. Senator William Cohen was a liberal Republican who later served as Secretary of Defense under Democratic President Bill Clinton; Senator Al D'Amato was exceptionally liberal on a number of social issues, who was endorsed for re-election by an LGBTQ advocacy group over a Democratic opponent (Nagourney 1998).

4.2.2 Neighbors in the Network

Geographical coalitions have been a longstanding feature in the Congress, with polarization between these coalitions intensifying in recent decades Abramowitz (2012). In this context, representing neighboring constituencies is often considered a firm basis for social connections between legislators (Arnold, Deen, and Patterson 2000; Fowler 2006; Kim, Barnett, and Kwon 2010).

In $\hat{S}$, more than 20% of the positive edges (122 out of 578) and none of the negative edges correspond to pairs of legislators representing the same or directly neighboring constituencies. We examine these connections in detail in Figure 2. Some communities have formed, including independent ones such as in the Northeast. To partition the larger interconnected network spanning the Midwest and the South, we apply Clauset-Newman-Moore greedy modularity maximization (Clauset, Newman, and Moore 2004), which aims to find the set of communities that produces the strongest division, i.e., dense intracommunity connections and sparse intercommunity connections. The detected communities are illustrated in Figure 2. Some communities show partisan tendencies, such as the one spanning Kansas, Missouri, Indiana, and Kentucky, forming sharp boundaries with the Midwestern and Southern Democrats. Many communities stretch across party lines, especially in the Northeast and Midwest.

4.2.3 Detecting Cosponsorship and Shared Committee Membership

Cosponsorship or shared committee membership are popular proxies for constructing social networks in the Congress (Porter et al. 2005, Fowler 2006, Kim, Barnett, and Kwon 2010, Battaglini, Sciabolazza, and Patacchini 2020). We examine whether $\hat{S}$ reflects either of these factors.

In his cosponsorship network, Fowler (2006) constructs the network edges based on direct sponsor-cosponsor relationships, ann indirect connections between cosponsors are not reflected on the edge weights. However, given that the average number of cosponsors per bill is 8.22 (excluding the sponsor) with a maximum of of 85 cosponsors, it seems unrealistic to assume that there is a meaningful social connection

$\hat{s}_{ij}$	Democrat	State	Republican	State	Com	Bills
0.809	Charles Robb	VA	*John Warner*	VA	1/1	181
0.779	Frank Lautenberg	NJ	*William Cohen*	ME	0/0	153
0.756	Bob Graham	FL	*Connie Mack*	FL	2/1	157
0.701	Joe Biden	DE	*Arlen Specter*	PA	1/0	107
0.627	James Exon	NE	*Larry Pressler*	SD	2/1	114
0.614	Tom Harkin	IA	*Larry Pressler*	SD	1/1	88
0.578	Richard Shelby	AL	*Trent Lott*	MS	1/1	129
0.546	Daniel Akaka	HI	*William Cohen*	ME	0/0	59
0.502	Max Baucus	MT	*Larry Pressler*	SD	1/1	77
0.495	Joe Lieberman	CT	*Al D'Amato*	NY	0/0	251

Table 3: Different-party senator pairs with the largest $\hat{s}_{ij}$ values. "Com" column indicates number of committees/subcommittees with intersecting membership, and "Bills" the number of cosponsored bills .

$\hat{s}_{ij}$	Senator	State	Senator	State
2.276	Richard Shelby	AL	Howell Heflin	AL
1.935	*John Heinz*	PA	*Arlen Specter*	PA
1.831	*Chuck Grassley*	IA	*Larry Pressler*	SD
1.738	Harry Reid	NV	Richard Bryan	NV
1.652	Sam Nunn	GA	Wyche Fowler	GA
1.441	David Pryor	AR	Dale Bumpers	AR
1.361	Sam Nunn	GA	Chuck Robb	VA
1.310	Robert Byrd	WV	Jay Rockefeller	WV
1.296	*Phil Gramm*	TX	*Don Nickles*	OK
1.288	*John Breaux*	LA	*J. Bennett Johnston*	LA

Table 4: Same-party senator pairs with the largest $\hat{s}_{ij}$ values. Italics indicate Republicans, and non-italics indicate Democrats.

Within-party	#	Sum	Across-party	#	Sum
John Heinz	15	3.39	*William Cohen*	18	1.86
Jake Garn	15	2.68	*Al D'Amato*	18	1.41
John Chafee	14	4.97	*Robert Packwood*	17	1.02
James Exon	14	4.14	*Larry Pressler*	17	2.1
Kent Conrad	14	4.02	*John Chafee*	17	0.24
Chuck Grassley	13	3.85	James Exon	17	1.07
Nancy Kassebaum	13	2.88	Howell Heflin	17	1.11
James Thurmond	12	3.46	Kent Conrad	17	0.24

Table 5: Most connected senators with the largest number of edges, with "Sum" indicating the sum of connected edge weights. Italics indicate Republicans, and non-italics indicate Democrats.

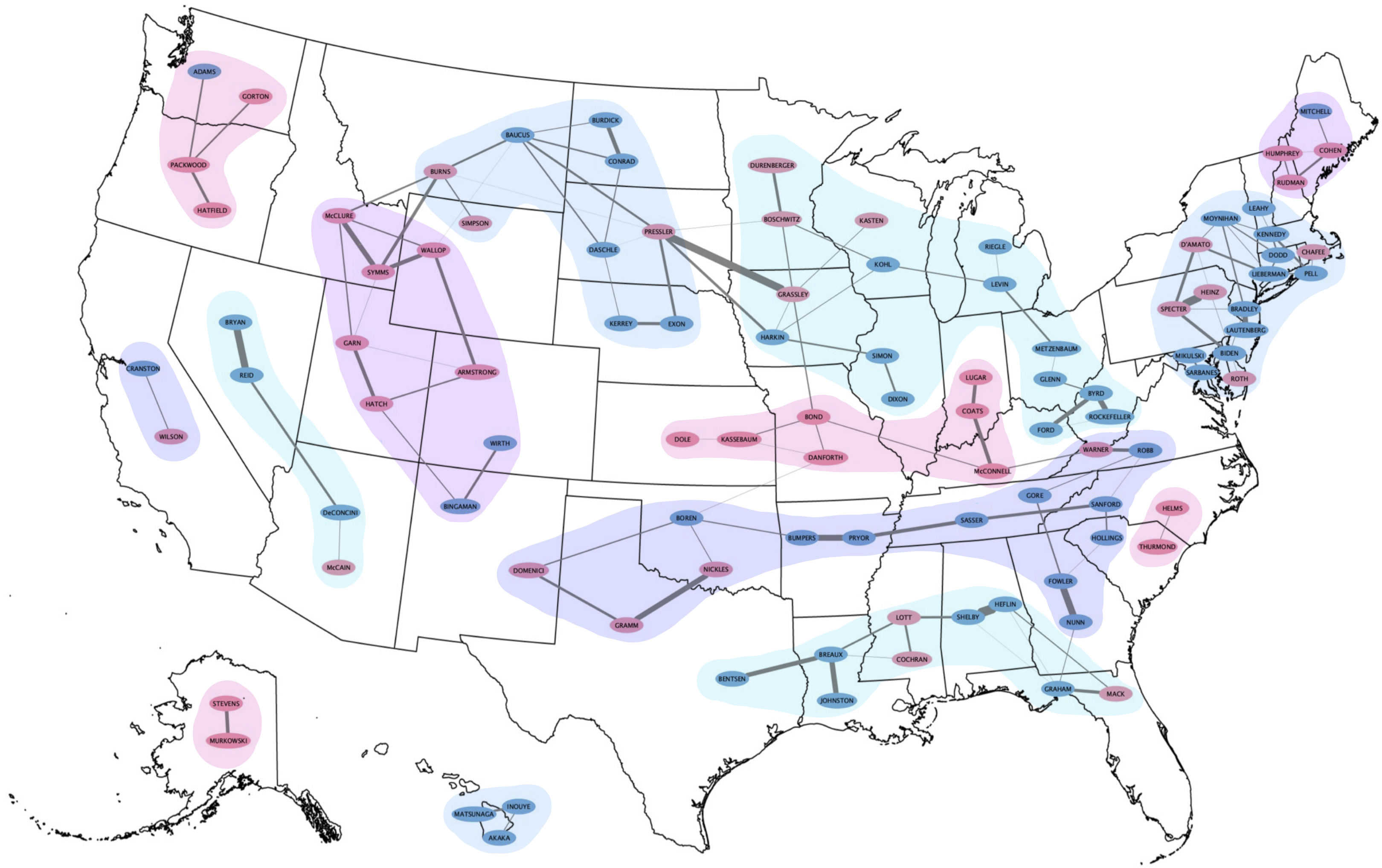

Figure 2: Sub-network of senators representing directly neighboring states, where each node represents a senator, placed on their state, and each edge a positive connection in $\hat{S}$. The color of the nodes represent the party affiliation of the senators, blue for Democrats and red for Republicans. The width of the edges correspond to the edge strengths, where wider edges indicate stronger connections. Colored groups indicate detected communities.

between sponsors and their cosponsors, but not between two cosponsors. Therefore, our cosponsorship network, $\mathcal{N}_{\text{cosp}}$, is constructed from both direct sponsor-cosponsor connections and indirect cosponsor-cosponsor connections. Therefore, the weight associated with the edge between two senators corresponds to the number of all co-signed bills, either as sponsors or cosponsors. The average number of cosponsored bills in $\mathcal{N}_{\text{cosp}}$ between any two senators is 80.66.

A comparison between $\hat{S}$ and $\mathcal{N}_{\text{cosp}}$ indicates that some edges of $\hat{S}$ correspond with higher cosponsorship, with over 66% of the edges in $\hat{S}$ correspond to cosponsorship levels above the median, and more than 22% correspond to the top 10% of cosponsorship levels. However, the correlation between the edge weights of $\hat{S}$ and $\mathcal{N}_{\text{cosp}}$ is a relatively low value of 0.18, so we conclude that there is only weak evidence of similarity between $\hat{S}$ and $\mathcal{N}_{\text{cosp}}$. This weak correlation may be explained by the fact that cosponsorship decisions are impacted by ideology and partisanship. This is supported by empirical evidence that partisan-ideological variables such as ADA scores, NOMINATE scores, and party affiliation are highly correlated with these decisions (Desposato, Kearney, and Crisp 2011).

For their House committee network, Porter et al. (2005) designate both the legislators and the committees as nodes, with edges connecting each legislator to the committees they are assigned to. Instead, our committee network $\mathcal{N}_{\text{comm}}$ contains only senators as nodes, where the edge weight corresponds to the number of committees or subcommittees the two senators share membership over. The resulting $\mathcal{N}_{\text{comm}}$ has 101 nodes and 2441 edges, where two senators on average have shared membership in 0.59 committees and 0.71 subcommittees.

We first observe that 56% (323/578) of the positive edges of $\hat{S}$ appear in $\mathcal{N}_{\text{comm}}$, while 47% of all possible senator pairs have some shared committee membership. Moreover, the correlation between the edge weights of $\hat{S}$ and $\mathcal{N}_{\text{comm}}$ is only 0.02. Overall, we conclude that there is no association between $\hat{S}$ and $\mathcal{N}_{\text{comm}}$. Similar to cosponsorship, committee activities may be better explained by party affiliation and the revealed preferences of senators, because senators are strategically appointed to committees by the party, and senators often seek appointments that are best aligned with their constituency interests.

To conclude, even though the cosponsorship and committee networks are often used as descriptors of social connections in the Congress, they show weak associations with $\hat{S}$. This suggests either that cosponsorship and committee networks are better explained by partisanship and revealed preferences of senators or that these connections in generality have weak legislative impact.

5 Discussion

Legislators' roll call decisions are influenced by a range of factors, including ideology, partisan strategies, and social connections. However, existing models of roll call behavior often focus on only one of these elements. Instead, we integrate the partisan-ideological and social approaches to roll call analysis through a fused latent factor and social network model. This model attributes the core patterns of voting to legislators' partisan-ideological positions through the latent factor component, while the remaining dependencies are interpreted as arising from social connections through the network component. Our approach provides a direct measure of social ties from roll call votes instead of relying on proxies, such as cosponsorship or shared committee membership.

An application of this model to the 101st Senate demonstrated that the estimated social network captures social connections that emerge from geographical proximity. The network also has low association with cosponsorship and shared committee membership networks. These findings suggest that geographical proximity leads to regional communities and ultimately shared roll call behavior, while cosponsorship and shared committee membership may follow partisan-ideological lines and/or have low legislative impact.

This applications serves as a "supervised" validation of the model, where we utilize well-studied data and existing analyses to confirm that the results are consistent with existing intuitions of roll call behavior. Going forward, we believe this model will be instrumental in "unsupervised" discoveries and measurements of social connections within both federal and local legislatures. We plan to apply this model to quantitatively analyze shifts in congressional social dynamics in the context of intensifying polarization.

Acknowledgements

For helpful comments, we thank Yunxiao Chen, Naoki Egami, Michael Sobel, Jacob Montgomery, Yiqing Xu, and poster session participants at the 40th Annual Meeting of the Society for Political Methodology.

References

Abramowitz, Alan I. (2012). "US Senate Elections in a Polarized Era". Vol. 27. CQ Press, p. 30.

Aldrich, John H and David W Rohde (2000). "The consequences of party organization in the House: The role of the majority and minority parties in conditional party government". *Polarized politics: Congress and the president in a partisan era*. Vol. 31. Citeseer, pp. 33–34.

Ansolabehere, Stephen, James M. Snyder, and Charles Stewart (2001). "The Effects of Party and Preferences on Congressional Roll-Call Voting". *Legislative Studies Quarterly* 26(4), pp. 533–572.

Arnold, Laura W., Rebecca E. Deen, and Samuel C. Patterson (2000). "Friendship and votes: The impact of interpersonal ties on legislative decision making". *State and Local Government Review* 32(2), pp. 142–147.

Arnold, R. Douglas (1990). *The Logic of Congressional Action*. Yale University Press.

Baker, Ross K. (1980). *Friend and Foe in the United States Senate*. Free Press.

Battaglini, Marco, Valerio Leone Sciabolazza, and Eleonora Patacchini (2020). "Effectiveness of Connected Legislators". *American Journal of Political Science* 64(4), pp. 739–756.

— (2023). "Abstentions and social networks in Congress". *The Journal of Politics* 85(2), pp. 581–592.

Boyd, Stephen, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein (2011). "Distributed optimization and statistical learning via the alternating direction method of multipliers". *Foundations and Trends in Machine learning* 3(1), pp. 1–122.

Bratton, Kathleen A. and Stella M. Rouse (2011). "Networks in the legislative arena: How group dynamics affect cosponsorship". *Legislative Studies Quarterly* 36(3), pp. 423–460.

Burkett, Tracy Lynn (1997). *Cosponsorship in the United States Senate: A network analysis of Senate communication and leadership, 1973-1990*. University of South Carolina.

Caldeira, Gregory A. and Samuel C. Patterson (1987). "Political friendship in the legislature". *The Journal of Politics* 49(4), pp. 953–975.

Canes-Wrone, Brandice, David W. Brady, and John F. Cogan (2002). "Out of Step, out of Office: Electoral Accountability and House Members' Voting". *The American Political Science Review* 96(1), pp. 127–140.

Caughey, Devin and Eric Schickler (2016). "Substance and change in congressional ideology: NOMINATE and its alternatives". *Studies in American Political Development* 30(2), pp. 128–146.

Chen, Yunxiao, Xiaoou Li, Jingchen Liu, and Zhiliang Ying (2016). "A fused latent and graphical model for multivariate binary data". *arXiv preprint arXiv:1606.08925*.

Clauset, Aaron, M. E. J. Newman, and Cristopher Moore (2004). "Finding community structure in very large networks". *Phys. Rev. E* 70 (6), p. 066111.

Clinton, Joshua, Simon Jackman, and Douglas Rivers (2004). "The statistical analysis of roll call data". *American Political Science Review* 98(2), pp. 355–370.

Desposato, Scott W., Matthew C. Kearney, and Brian F. Crisp (2011). "Using Cosponsorship to Estimate Ideal Points". *Legislative Studies Quarterly* 36(4), pp. 531–565.

Downs, Anthony (1957). *An Economic Theory of Democracy*. Harper & Brothers.

Everson, Phil, Rick Valelly, Arjun Vishwanath, and Jim Wiseman (2016). "NOMINATE and American political development: a primer". *Studies in American Political Development* 30(2), pp. 97–115.

Fong, Christian (2020). "Expertise, networks, and interpersonal influence in congress". *The Journal of Politics* 82(1), pp. 269–284.

Fowler, James H. (2006). "Connecting the Congress: A study of cosponsorship networks". *Political analysis* 14(4), pp. 456–487.

Heckman, James J. and James M. Snyder (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.

Kim, Jang Hyun, George A. Barnett, and K. Hazel Kwon (2010). “The influence of social networks on the US Senate roll-call voting”. *International Journal of E-Politics (IJEP)* 1(4), pp. 22–41.

Kirkland, Justin H (2011). “The relational determinants of legislative outcomes: Strong and weak ties between legislators”. *The Journal of Politics* 73(3), pp. 887–898.

Kirkland, Justin H. and Justin H. Gross (2014). “Measurement and theory in legislative networks: The evolving topology of Congressional collaboration”. *Social Networks* 36. Special Issue on Political Networks, pp. 97–109.

Klein, Ezra (2009). *Arlen Specter and Joe Biden: An amtrak love story.*

Krehbiel, Keith (1993). “Where’s the Party?” *British Journal of Political Science* 23(2), pp. 235–266.

Lee, Frances E. (2009). *Beyond Ideology: Politics, Principles, and Partisanship in the U. S. Senate.* University of Chicago Press.

Mayhew, David R. (1974). *Congress: The Electoral Connection.* Yale studies in political science 26. Yale University Press.

McCarty, Nolan, Keith T Poole, and Howard Rosenthal (2016). *Polarized America: The dance of ideology and unequal riches.* MIT Press.

Miller, Warren E. and Donald E. Stokes (1963). “Constituency Influence in Congress”. *American Political Science Review* 57(1), pp. 45–56.

Minozzi, William and Craig Volden (2013). “Who Heeds the Call of the Party in Congress?” *The Journal of Politics* 75(3), pp. 787–802.

Nagourney, Adam (1998). *D’Amato Wins Endorsement Of Gay Group.*

Pianin, Eric (1991). *House minority leader Michel enjoys political revival of sorts.*

Poole, Keith T. and Howard Rosenthal (1985). “A Spatial Model for Legislative Roll Call Analysis”. *American Journal of Political Science* 29(2), pp. 357–384.

Poole, Keith T. and Howard Rosenthal (2000). *Congress: A political-economic history of roll call voting.* Oxford University Press, USA.

Porter, Mason A., Peter J. Mucha, Mark E.J. Newman, and Casey M. Warmbrand (2005). "A network analysis of committees in the US House of Representatives". *Proceedings of the National Academy of Sciences* 102(20), pp. 7057–7062.

Ramey, Adam J., Jonathan D. Klingler, and Gary E. Hollibaugh Jr. (2017). *More Than a Feeling: Personality, Polarization, and the Transformation of the US Congress.* University of Chicago Press.

Rohde, David W. (1991). *Parties and leaders in the postreform House.* University of Chicago Press.

Sinclair, Barbara (2016). *Unorthodox lawmaking: New legislative processes in the US Congress.* CQ Press.

Snyder, James M. and Tim Groseclose (2000). "Estimating Party Influence in Congressional Roll-Call Voting". *American Journal of Political Science* 44(2), pp. 193–211.

Stauffer, Dietrich (2008). "Social applications of two-dimensional Ising models". *American Journal of Physics* 76(4), pp. 470–473.

Wawro, Gregory J. and Eric Schickler (2007). *Filibuster: Obstruction and lawmaking in the US Senate.* Vol. 134. Princeton University Press.

Wright, John and Yi Ma (2022). *High-dimensional data analysis with low-dimensional models: Principles, computation, and applications.* Cambridge University Press.