A Statistical Model of Bipartite Networks: Application to Cosponsorship in the United States Senate.*

Adeline Lo† Santiago Olivella‡ Kosuke Imai§

July 2025

*The methods described in this paper can be implemented via the open-source statistical software, NetMix, available at https://CRAN.R-project.org/package=NetMix. The authors are grateful for comments from Alison Craig, Skyler Cranmer, Sarah Shugars and the participants of the Harvard IQSS Applied Statistics seminar.

†Associate Professor of Political Science, UW-Madison. Email: aylo@wisc.edu, URL:https://www.loadeline.com/

‡Associate Professor of Political Science, UNC-Chapel Hill. Email: olivella@unc.edu

§Professor, Department of Government and Department of Statistics, Harvard University. 1737 Cambridge Street, Institute for Quantitative Social Science, Cambridge 02138. Email: imai@harvard.edu, URL:https://imai.fas.harvard.edu/

Contents

S.1 Additional methodological details

S.1.1 Plate Diagram of the Proposed Model

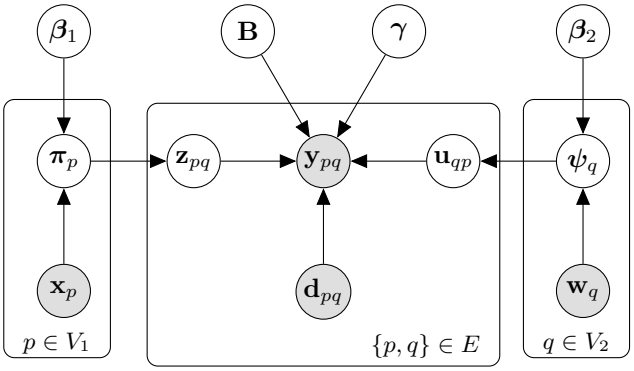


Figure S.1: **Plate diagram of the proposed model**. Observed data represented as shaded nodes; hyperparameters presented as nodes outside plates.

S.1.2 Motivation for the proposed model components

The model components in Figure S.1 are most easily justified when considering the case of cosponsorship, legislative productivity, and collaboration . Any modeling tool that hopes to produce plausible answers to the puzzle of productivity and collaboration in times of nominal partisan division should allow researchers to explore the heterogeneity of both legislator and bill attributes, as allowed by the $\mathbf{x_p}$ and $\mathbf{w_q}$ terms in our model. Much as senators might differentially decide to cosponsor based on their personal characteristics, so too might legislation attract cosponsorship based on its different content and the specific context in which it is introduced. A useful model of cosponsorships thus allows the politics of coalition formation around certain types of legislation to account for this heterogeneity.

Furthermore, it is important for such a model to capture how the interaction of different bill and senator types can lead to cosponsorship collaborations beyond what would be expected from understanding the coalitional dynamics that drive much of legislative politics, as allowed by the $\mathbf{d_{pq}}$ term in Figure S.1. The extant literature has identified two substantial predictors of cosponsorship decision that are defined at the senator-bill dyad level, and that result in the kind of non-coalitional homophily and heterophily that is common in social networks of different kinds.

First, individual senators often trade favors with one another, such that cosponsorship may result from quid pro quo behavior and norms of individual reciprocity — *you cosponsored my bill before, I'll cosponsor yours now* (Brandenberger, 2018; Harbridge-Yong, Volden, and Wiseman, 2023). Similarly, scholars have increasingly stressed the role of Senate committees in forming support around legislation, finding evidence that sitting in a committee involved in the life cycle of a bill can affect a legislator's support for it (Porter et al., 2005; Cirone and Van Coppenolle, 2018). Accordingly, a model that aims to capture the full set of forces behind a bipartite network such as that formed by cosponsorships should account for these kinds of naturally dyadic features that complement group-based drivers of edge formation.

S.1.3 Details of the Estimation Algorithm

To approximate the collapsed posterior proportional to Equation (6), we first define a factorized distribution of the joint dyad-specific latent group membership variables (i.e. $\mathbf{Z}$ and $\mathbf{U}$) as follows:

$$m(\mathbf{Z}, \mathbf{U} \mid \mathbf{\Phi}) = \prod_{p,q \in V_1 \times V_2} m(z_{pq}, u_{pq} \mid \boldsymbol{\phi}_{pq}) \tag{S.1}$$

where $\mathbf{\Phi} = \{\boldsymbol{\phi}_{pq}\}_{p,q \in V_1 \times V_2}$ are sum-to-one, $(K_1 \times K_2)$-dimensional variational parameters.

The goal of variational inference is to find, in the space of functions of the form given by Equation (S.1), one that closely approximates (in KL divergence terms) the target posterior. This is equivalent to maximizing the evidence lower bound $\mathcal{L}(\mathbf{\Phi})$ to Equation (6)with respect to vectors $\boldsymbol{\phi}_{pq}$:

$$\begin{aligned} \hat{\boldsymbol{\phi}}_{pq} &= \underset{\boldsymbol{\phi}_{pq}}{\operatorname{argmax}} \underbrace{\mathbb{E}_m \left[\log f(\mathbf{Y}, \mathbf{Z}, \mathbf{U}, \mid \mathbf{B}, \boldsymbol{\beta}, \boldsymbol{\gamma})\right] - \mathbb{E}_m \left[\log m(\mathbf{Z}, \mathbf{U} \mid \mathbf{\Phi})\right]}_{\mathcal{L}(\mathbf{\Phi})} \\ &\propto\!\!\!\!\sim \{(\alpha_{pg} + C'_{pg})(\alpha_{qg} + C'_{qh})(\theta^{y_{pq}}_{pq,g,h}(1 - \theta_{pq,g,h})^{1-y_{pq}})\}_{g,h \in K_1 \times K_2} \end{aligned} \tag{S.2}$$

where $C'_{pg} = \sum_{q' \in V_2^{-q}} \sum_{h'=1}^{K_2} \phi_{pq',g,h'}$ is the expected value of the marginal count C_{pg} under the variational distribution (and similarly for C'_{qh}). The approximation in the last line results from using a zeroth-order Taylor series expansion of the expectation in place of calculating the computationally expensive integral over the Poisson-Binomial distribution of the count statistics.

In addition to finding the posterior over mixed-membership vectors, we take an empirical Bayes approach and maximize the lower bound $\mathcal{L}(\mathbf{\Phi})$ to obtain values of relevant hyper-parameters $\mathbf{B}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$.

After appropriate initialization (see below), then, the full lower bound optimization proceeds iteratively by first updating the variational parameters according to Equation (S.2) (the E-step), and maximizing $\mathcal{L}(\mathbf{\Phi})$ w.r.t. to the hyper-parameters, holding $\boldsymbol{\phi}_{pq}$ (and derived global counts C_{pg} and C_{qh}) constant at their most recent value (the M-step), until the change in the lower bound is below a user-specified tolerance. As there are no closed-form solution for these optimal values, we rely on a numerical optimization routine (required gradients are available below). In what follows, we provide details on those steps.

S.1.3.1 E-step

E-step: Z and U

Variational parameters $\boldsymbol{\phi}_{pq}$ are updated by restricting Equation (S.2) to terms that depend only on $\mathbf{z}_{pq}$ and $\mathbf{u}_{pq}$ and taking the logarithm of the resulting expression,

$$\begin{aligned}
&\log P(\mathbf{Y}, \mathbf{Z}, \mathbf{U} \mid \mathbf{B}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma}, \mathbf{X}_1, \mathbf{X}_2, \mathbf{D}) \\
&= z_{pq,g} \sum_{h=1}^{K_2} u_{qq,h} \left\{ Y_{pq} \log(\theta_{pqgh}) + (1 - Y_{pq}) \log(1 - \theta_{pqgh}) \right\} + \log \Gamma(\alpha_{pg} + C_{pg}) + \text{const.}
\end{aligned}$$

Note that $C_{pg} = C'_{pg} + \mathbb{1}(z_{pq,g} = g)$ and that, for $x \in \{0, 1\}$, $\Gamma(y+x) = y^x \Gamma(y)$. As $\mathbb{1}(z_{pq,g} = g) \in \{0, 1\}$, we can re-express $\log \Gamma(\alpha_{pg} + C_{pg}) = z_{pq,g} \log(\alpha_{pg} + C'_{pg}) + \log \Gamma(\alpha_{pg} + C'_{pg})$ and simplify the expression:

$$z_{pq,g} \sum_{h=1}^{K_2} u_{qp,h} \left\{ Y_{pq} \log(\theta_{pqgh}) + (1 - Y_{pq}) \log(1 - \theta_{pqgh}) \right\} + z_{pq,g} \log \left(\alpha_{pg} + C'_{pg} \right) + \text{const.}$$

Then take the expectation under the variational distribution $\widetilde{Q}$:

$$\begin{aligned}
&\mathbb{E}_{\widetilde{Q}} \{ \log P(\mathbf{Y}, \mathbf{Z}, \mathbf{U} \mid \mathbf{B}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma}, \mathbf{D}, \mathbf{X}_1, \mathbf{X}_2) \} \\
&= z_{pq,g} \sum_{h=1}^{K_2} \mathbb{E}_{\widetilde{Q}}(u_{qp,g}) \big(Y_{pqt} \log(\theta_{pqgh}) + (1 - Y_{pq}) \log(1 - \theta_{pqgh}) \big) + z_{pq,g} \, \mathbb{E}_{\widetilde{Q}} \left\{ \log \left(\alpha_{pg} + C'_{pg} \right) \right\} + \text{const.}
\end{aligned}$$

The exponential of this corresponds to the (unnormalized) parameter vector of a multinomial distribution $\widetilde{Q}(\mathbf{z}_{pq} \mid \boldsymbol{\phi}_{pq})$.

The update for $\mathbf{u}_{qp}$ is similarly derived. Restrict Equation (S.2) to terms that depend only on $\mathbf{u}_{qp}$ (for specific p, q nodes in V) and taking the logarithm of the resulting expression,

$$\log P(\mathbf{Y}, \mathbf{Z}, \mathbf{U} \mid \mathbf{B}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma}, \mathbf{X}_1, \mathbf{X}_2, \mathbf{D})$$

$$= u_{qp,h} \sum_{g=1}^{K_1} z_{pq,g} \left\{ Y_{pq} \log(\theta_{pqgh}) + (1 - Y_{pq}) \log(1 - \theta_{pqgh}) \right\} + \log \Gamma(\alpha_{qh} + C_{qh}) + \text{const.}$$

Re-express $\log \Gamma(\alpha_{qh} + C_{qh}) = u_{qp,h} \log(\alpha_{qh} + C'_{qh}) + \log \Gamma(\alpha_{qh} + C'_{qh})$ and simplify the expression:

$$u_{qp,h} \sum_{g=1}^{K_1} z_{pq,g} \left\{ Y_{pq} \log(\theta_{pqgh}) + (1 - Y_{pq}) \log(1 - \theta_{pqgh}) \right\} + u_{qp,h} \log \left(\alpha_{2qh} + C'_{qh} \right) + \text{const.}$$

Take the expectation under the variational distribution $\widetilde{Q}$:

$$\begin{aligned}
&\mathbb{E}_{\widetilde{Q}} \{ \log P(\mathbf{Y}, \mathbf{Z}, \mathbf{U} \mid \mathbf{B}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma}, \mathbf{D}, \mathbf{X}_1, \mathbf{X}_2) \} \\
&= u_{qp,h} \sum_{g=1}^{K_1} \mathbb{E}_{\widetilde{Q}}(z_{pq,g}) \big(Y_{pq} \log(\theta_{pqgh}) + (1 - Y_{pq}) \log(1 - \theta_{pqgh}) \big) + u_{qp,h} \;\; \mathbb{E}_{\widetilde{Q}} \left\{ \log \left(\alpha_{2qh} + C'_{qh} \right) \right\} + \text{const.}
\end{aligned}$$

S.1.3.2 M-step

Lower Bound

Expression for the lower bound,

$$\begin{aligned}
\mathcal{L}(\tilde{Q}) \;=\; & \mathbb{E}_{\widetilde{Q}}[\log P(\mathbf{Y}, \mathbf{Z}, \mathbf{U} \mid \mathbf{B}, \boldsymbol{\gamma}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \mathbf{X}_1, \mathbf{X}_2, \mathbf{D})] - \mathbb{E}_{\widetilde{Q}}[\log \widetilde{Q}(\mathbf{Z}, \mathbf{U} \mid \boldsymbol{\Phi})] \\
= & \sum_{p \in V_1} \Big[\log \Gamma(\xi_p) - \log \Gamma(\xi_p + N_2) \Big] + \sum_{q \in V_2} \Big[\log \Gamma(\xi_q) - \log \Gamma(\xi_q + N_1) \Big] \\
& + \sum_{p \in V_1} \sum_{g=1}^{K_1} \Big[\mathbb{E}[\log \Gamma(\alpha_{pg} + C_{pg})] - \log \Gamma(\alpha_{pg}) \Big] + \sum_{q \in V_2} \sum_{h=1}^{K_2} \Big[\mathbb{E}[\log \Gamma(\alpha_{qh} + C_{qh})] - \log \Gamma(\alpha_{qh}) \Big] \\
& + \sum_{(p,q) \in V_1 \times V_2} \sum_{g=1}^{K_1} \sum_{h=1}^{K_2} \phi_{pq,g} \phi_{pq,h} \left\{ Y_{pq} \log \theta_{pqgh} + (1 - Y_{pq}) \log(1 - \theta_{pqgh}) \right\} \\
& - \sum_{g,h=1}^{K} \frac{(B_{gh} - \mu_{gh})^2}{2\sigma_{gh}^2} - \sum_{j=1}^{J_d} \frac{(\gamma_j - \mu_\gamma)^2}{2\sigma_\gamma^2} - \sum_{g=1}^{K_1} \sum_{j=1}^{J_{1x}} \frac{(\beta_{1gj} - \mu_{\beta_1})^2}{2\sigma_{\beta_1}^2} - \sum_{h=1}^{K_2} \sum_{j=1}^{J_{2x}} \frac{(\beta_{2hj} - \mu_{\beta_2})^2}{2\sigma_{\beta_2}^2} \\
& - \sum_{(p,q) \in V_t} \sum_{g=1}^{K_1} \sum_{h=1}^{K_2} \{ \phi_{pq,g} \log \phi_{pq,g} - \phi_{qp,h} \log(\phi_{qp,h}) \}
\end{aligned}$$

M-step 1: update for B

Restricting the lower bound to terms that contain B_{gh} (blockmodel), we obtain

$$\mathcal{L}(\widetilde{Q}) \;=\; \sum_{p,q \in E_t} \sum_{g=1}^{K_1} \sum_{h=1}^{K_2} \phi_{pq,g} \phi_{qp,h} \{ Y_{pq} \log \theta_{pqgh} + (1 - Y_{pq}) \log(1 - \theta_{pqgh}) \}$$

$$- \sum_{g,h=1}^{K} \frac{(B_{gh} - \mu_{gh})^2}{2\sigma_{gh}^2} + \text{const.}$$

Optimize this lower bound with respect to $\mathbf{B}_{gh}$ using a gradient-based numerical optimization method. The corresponding gradient is:

$$\frac{\partial \mathcal{L}_{B_{gh}}}{\partial B_{gh}} = \sum_{p,q \in V_1 \times V_2} \phi_{pq,g}\phi_{qp,h} \left(Y_{pq} - \theta_{pqgh}\right) - \frac{B_{gh} - \mu_{B_{gh}}}{\sigma_{B_{gh}}^2}$$

M-step 2: update for $\boldsymbol{\gamma}$

Restricting the lower bound to terms containing $\boldsymbol{\gamma}$ (dyadic coefficients), and recalling that $\theta_{pqtgh} = [1 + \exp(-B_{gh} - \mathbf{d}_{pqt}\boldsymbol{\gamma})]^{-1}$, then:

$$\begin{aligned}\mathcal{L}(\tilde{Q}) = & \sum_{p,q \in V_1 \times V_2} \sum_{g=1}^{K_1}\sum_{h=1}^{K_2} \phi_{pq,g}\phi_{qp,h} \left\{Y_{pq} \log \theta_{pqgh} + (1 - Y_{pq}) \log(1 - \theta_{pqgh})\right\} \\ & - \sum_{j}^{J_d} \frac{(\gamma_j - \mu_\gamma)^2}{2\sigma_\gamma^2} + \text{const.}\end{aligned}$$

To optimize this expression w.r.t. γ_j (the jth element of the $\boldsymbol{\gamma}$ vector), we again use a numerical optimization algorithm based on the following gradient,

$$\frac{\partial \mathcal{L}(\widetilde{Q})}{\gamma_j} = \sum_{p,q \in V_1 \times V_2} \sum_{g=1}^{K_1}\sum_{h=1}^{K_2} \phi_{pq,g}\phi_{qp,h} d_{pqj} \left(Y_{pq} - \theta_{pqgh}\right) - \frac{\gamma_j - \mu_\gamma}{\sigma_\gamma^2}$$

M-step 3: update for $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$

Let $\alpha_{pg} = \exp\left(\mathbf{x}_1^\top \boldsymbol{\beta}_{1g}\right)$, $\xi_p = \sum_{g=1}^{K_1} \alpha_{pg}$, $\alpha_{qh} = \exp\left(\mathbf{x}_{2q}^\top \boldsymbol{\beta}_{2h}\right)$, and $\xi_q = \sum_{h=1}^{K_2} \alpha_{qh}$. To find the optimal value of $\boldsymbol{\beta}_{1g}$, roll all terms not involving the coefficient vector into a constant:

$$\begin{aligned}\mathcal{L}(\widetilde{Q}) = & \sum_{p \in V_1} \left[\log \Gamma(\xi_{1p}) - \log \Gamma(\xi_p + N_2)\right] \\ & + \sum_{p \in V_1} \sum_{g=1}^{K_1} \left[\mathbb{E}_{\widetilde{Q}_2}[\log \Gamma(\alpha_{pg} + C_{pg})] - \log \Gamma(\alpha_{pg})\right] \\ & - \sum_{g=1}^{K_1} \sum_{j=1}^{J_{1x}} \frac{(\beta_{1gj} - \mu_{\beta_1})^2}{2\sigma_{\beta_1}^2} + \text{const.}\end{aligned}$$

No closed form solution exists for an optimum w.r.t. β_{1gj}, but a gradient-based algorithm can be implemented to maximize the above. The corresponding gradient w.r.t. each element of $\boldsymbol{\beta}_{1g}$ is:

$$\begin{aligned}\frac{\partial \mathcal{L}(\tilde{Q})}{\partial \beta_{1gj}} \;=\; & \sum_{p \in V_1} \alpha_{pg} x_{1pj} \Big(\mathbb{E}_{\tilde{Q}_2}[\breve{\psi}(\alpha_{pg} + C_{pg}) - \breve{\psi}(\alpha_{pg})] \\ & + \Big[\breve{\psi}(\xi_{1p}) - \breve{\psi}(\xi_{1p} + N_1) \Big] \Big) \\ & - \frac{\beta_{1gj} - \mu_{\beta_1}}{\sigma^2_{\beta_1}}\end{aligned}$$

where $\breve{\psi}(\cdot)$ is the digamma function. Again, we can approximate expectations of non-linear functions of random variables using a zeroth-order Taylor series expansion. The M-step for the regression coefficients of the second family is similarly defined.

S.1.4 Stochastic variational inference

On the tth iteration, our algorithm completes the following steps:

1. Sample a subset of dyads $E^t \subset E$, with corresponding sets of vertices $V_1^t = \{p : p, q \in E^t\}$ and $V_2^t = \{q : p, q \in E^t\}$.

2. Update all $\boldsymbol{\phi}_{pq:p,q \in E_t}$ according to Equation (S.2), and compute a set of intermediate global count statistics (after normalization),

$$\widehat{C}_{pg} = \frac{N_2}{|V_2^t|} \sum_{q' \in V_2^t} \sum_{h'=1}^{K_2} \phi_{pq',g,h'}; \quad \widehat{C}_{qh} = \frac{N_1}{|V_1^t|} \sum_{p' \in V_1^t} \sum_{g'=1}^{K_1} \phi_{p'q,g',h}$$

   weighted to match the amount of information contained in the original network.

3. Update global count statistics matrices using an online average that follows an appropriately decreasing step-size schedule:

$$\mathbf{C}_p^{(t)} = (1 - \rho_{p,t}) \mathbf{C}_p^{(t-1)} + \rho_{p,t} \widehat{\mathbf{C}}_p; \quad \mathbf{C}_q^{(t)} = (1 - \rho_{q,t}) \mathbf{C}_q^{(t-1)} + \rho_{q,t} \widehat{\mathbf{C}}_q$$

   where step-size $\rho_{p,t} = (\tau + t)^{\kappa}$ such that $\tau > 0$ and $\kappa \in (0.5, 1]$.

4. Update values of hyper-parameters $\Lambda = \{\boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{B}\}$ by taking an "online" step in the direction of the (noisy) Euclidean gradient of $\mathcal{L}(\boldsymbol{\Phi})$ w.r.t Λ:

$$\lambda^{(t)} = \lambda^{(t-1)} + \rho_{\lambda,t} \nabla_{\Lambda} \mathcal{L}^{(t)}(\boldsymbol{\Phi})$$

with appropriate gradients given in Appendix.

Although different dyad sampling heuristics used for Step 1 can result in unbiased gradient estimates (see Gopalan and Blei, 2013, for a few examples), we follow the scheme proposed by Dulac, Gaussier, and Largeron, 2020, which is both simple to implement and has been found to work well in sparse settings. The procedure is:

1. Sample node i in V uniformly at random.

2. Form a set $s_1 = \{i, j : y_{ij} = 1, \forall j \in V\}$ (i.e., set of all connected dyads involving i). Form M sets $s_0^m = \{i, j : y_{ij} = 0, \exists j \in V\}$ (i.e., a set of some disconnected dyads involving i), where each set is of equal cardinality, and the disconnected dyads are sampled uniformly at random and with replacement.

3. Sample, with equal probability, either s_1 or any of the s_0^m sets. This set of dyads constitutes a subnetwork.

In our application, we set $M = 10$ and set $|s_0^m|$ be $1/M$ times the number of non-links between i and every other vertex in the network.

After the algorithm converges, we can recover the mixed-membership vectors by computing their posterior predictive expectations:

$$\widehat{\pi}_{pg} = \frac{C_{pg} + \alpha_{pg}}{N_2 + \sum_{g'=1}^{K_1} \alpha_{pg'}}, \quad \widehat{\psi}_{qh} = \frac{C_{qh} + \alpha_{qh}}{N_1 + \sum_{h'=1}^{K_2} \alpha_{qh'}}$$

S.1.5 Initial values for ϕ and ψ

Implementation of the model requires defining good starting values for the mixed-membership vectors. While spectral clustering methods offer good starting values for π and ψ in the unipartite setting, applying it to non-square affiliation matrices poses interesting challenges. To produce high-quality initial values in a viable amount of time we rely on the co-clustering approach of (Govaert and Nadif, 2003), which estimates a simpler, single-membership SBM using a fast EM algorithm.

S.1.6 Standard error computation

We obtain measures of uncertainty around regression coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ by evaluating the curvature of the lower bound at the estimated optimal values for these hyper-parameters. When considering terms that involve these hyper-parameters, the lower bound reduces to the expected value of the log-posterior taken with respect to the variational distribution $\widetilde{Q}$. Thus, evaluating the Hessian of the lower bound (and the corresponding covariance matrix of the hyper-parameters) requires evaluating that expectation. Details of the required Hessian are below.

S.1.6.1 Hessian for γ

Restricted to terms that involve $\boldsymbol{\gamma}$, the typical element of the required Hessian is given by

$$\frac{\partial^2 \mathcal{L}(\widetilde{Q})}{\partial \gamma_j \partial \gamma_{j\prime}} = \sum_{p,q \in V_1 \times V_2} -d_{pqj} d_{pqj\prime} \left[\bar{\theta}_{pq}(1-\bar{\theta}_{pq})\right] - \sigma_{\gamma}^{-2}\delta_{jj\prime}$$

where $\delta_{jj\prime}$ is the Kronecker delta function, and the term

$$\bar{\theta}_{pq} = \mathbb{E}_{\widetilde{Q}}[\theta_{pq}] = \hat{\boldsymbol{\phi}}_{pq}^{\top} \hat{\mathbf{B}} \hat{\boldsymbol{\phi}}_{qp} + \mathbf{d}_{pq}^{\top} \hat{\boldsymbol{\gamma}}$$

is a closed-form solution to the expectation over the variational distribution of the model's parameters.

S.1.6.2 Hessian for β_1 and β_2

In turn, and focusing on Family 1 coefficients, we can characterize the Hessian of the lower bound w.r.t. $\boldsymbol{\beta}_1$ with

$$\begin{aligned}\frac{\partial^2 \mathcal{L}(\widetilde{Q})}{\partial \beta_{gj} \partial \beta_{gj\prime}} = \sum_{p \in V_1} x_{pj} x_{pj\prime} \alpha_{pg} \Bigg(& \breve{\psi}(\xi_p) - \breve{\psi}(\alpha_{pg}) + \mathbb{E}_{\widetilde{Q}}[\breve{\psi}(\alpha_{pg} + C_{pg})] - \breve{\psi}(\xi_p + N_2) \\ & + \alpha_{pg} \left(\breve{\psi}_1(\xi_{pg}) - \breve{\psi}_1(\alpha_{pg}) + \mathbb{E}_{\widetilde{Q}}[\breve{\psi}_1(\alpha_{pq} + C_{pq})] - \breve{\psi}_1(\xi_p + N_2) \right) \Bigg) \\ & - \sigma_{\beta_1}^{-2} \delta_{jj\prime}\end{aligned}$$

for coefficients in the same group g, and

$$\frac{\partial^2 \mathcal{L}(\widetilde{Q})}{\partial \beta_{gj} \partial \beta_{hj\prime}} = \sum_{p \in V_1} x_{pj} x_{pj\prime} \alpha_{pg} \alpha_{ph} \left(\breve{\psi}_1(\xi_p) - \breve{\psi}_1(\xi_p + N_2) \right)$$

for coefficients associated with different latent groups g and h. As before, we use $\breve{\psi}(\cdot)$ to denote the digamma function, and $\breve{\psi}_1(\cdot)$ to denote the trigamma function.

Unlike the Hessian for $\boldsymbol{\gamma}$, there are no closed-form solutions for the expectations involved in these expressions. To approximate them, we take S samples from the Poisson-Binomial distribution of C_{pg}, $C_{pg}^{(s)}, s \in 1, \ldots, S$, and let

$$\mathbb{E}_{\widetilde{Q}}[\breve{\psi}(\alpha_{pg} + C_{pg})] \approx \frac{1}{S}\sum_s \breve{\psi}(\alpha_{pq} + C_{pq}^{(s)}); \quad \mathbb{E}_{\widetilde{Q}}[\breve{\psi}_1(\alpha_{pq} + C_{pq})] \approx \frac{1}{S}\sum_s \breve{\psi}_1(\alpha_{pq} + C_{pq}^{(s)})$$

The Hessian for the coefficients associated with Family 2, $\boldsymbol{\beta}_2$, is similarly approximated.

S.2 Simulation results

Setup. We simulate bipartite networks with unbalanced numbers of Senator and Bills nodes under 6 different scenarios, defined by overall network size and difficulty of the mixed-membership learning problem. More specifically, we define small (i.e. 300 total nodes) and large (i.e. 3000 total nodes) networks, each with twice as many Bill nodes as there are Senator nodes. In all instances, we define the edge-generating process according to our model, using $K_1 = K_2 = 2$ latent groups for each of the node types, a single monadic predictor drawn independently from $N(0, 1.5)$, and a single irrelevant dyadic predictor drawn from a standard Normal distribution.

Scenario Difficulty	**Blockmodel**	**Monadic Coefficients**
Easy	$\begin{bmatrix} 0.85 & 0.01 \\ 0.01 & 0.99 \end{bmatrix}$	$\begin{bmatrix} -4.50 & -4.50 \\ 0.00 & 0.00 \end{bmatrix}$
Medium	$\begin{bmatrix} 0.65 & 0.35 \\ 0.20 & 0.75 \end{bmatrix}$	$\begin{bmatrix} 0.05 & 0.75 \\ -0.75 & -1.00 \end{bmatrix}$
Hard	$\begin{bmatrix} 0.65 & 0.40 \\ 0.50 & 0.45 \end{bmatrix}$	$\begin{bmatrix} 0.00 & 0.00 \\ -0.75 & -1.00 \end{bmatrix}$

Table S.1: Simulation scenarios, defining various levels of estimation difficulty

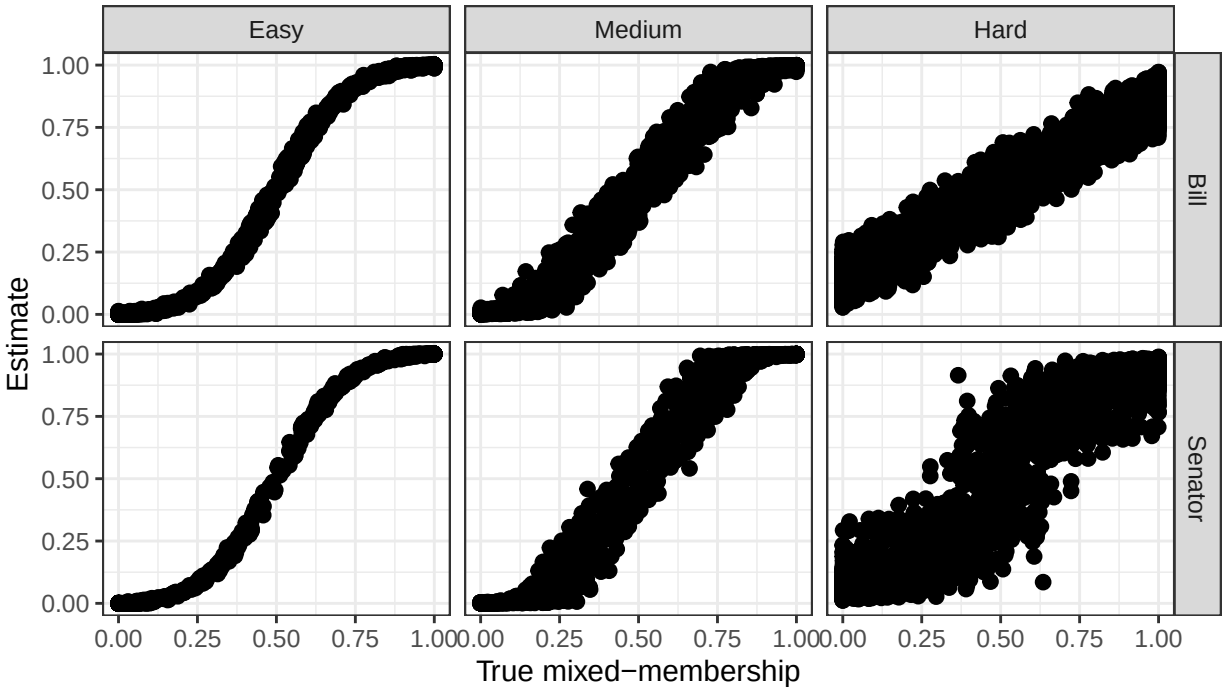


Figure S.2: **Mixed-Membership Recovery**: Estimated and true mixed-membership vectors under the easy, medium, and hard estimation scenarios. In all instances, recovery is excellent.

To simulate different levels of estimation difficulty, we vary both the blockmodel and the coefficients associated with the mixed-membership vectors, which are set to be equal across the two node types. In the "easy" scenario, memberships are barely mixed, and there is a clear difference in edge probabilities between different groups of the different node types. In contrast, the "hard" scenario is such that all nodes have a roughly equal probability of instantiating each block, and there is little difference in the probabilities of forming edges between blocks, as given by the blockmodel. The "medium" scenario offers a more realistic, in-between estimation problem. The specific values for scenarios are given in Table S.1.

Results. We begin by evaluating the accuracy of mixed-membership estimation by comparing true and estimated mixed-membership vectors (after re-labeling the latter to match the known, simulated group labels using the Hungarian algorithm). Correlations across node types and difficulty scenarios are demonstrated in Figure S.2. Overall, our model retrieves these mixed-membership vectors with a high degree of accuracy — even in regimes in which block memberships play a small role in the generation of edges, and regardless of whether there is an asymmetry in the number of nodes in each family.

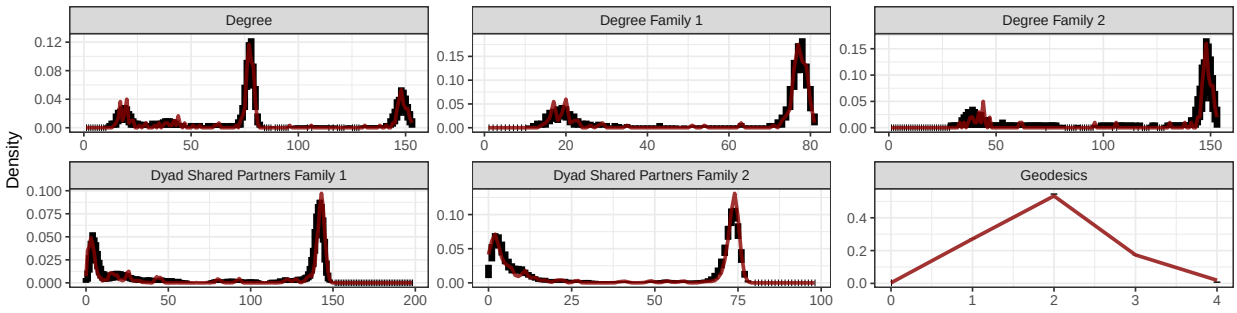


Figure S.3: **Posterior Predictive Goodness-Of-Fit**: The figure shows, for a randomly chosen simulated scenario, the extent to which the model can recover structural features of the observed network. The solid red line traces the observed distribution of the different network motifs; solid black rectangles show the central 90% distribution of values observed across 100 network replicates, obtained from the estimated model posterior. A good fit is indicated by lines that always fall within black regions.

Next, we evaluate the accuracy of estimated node-level and dyadic coefficients by simulating derived quantities of interest based on them, and compare these simulated quantities to their true counterparts, as one would when conducting a goodness-of-fit analysis based on posterior predictive distributions. As is typical in network modeling, these derived quantities are structural features of the network (e.g. Hunter, Goodreau, and Handcock, 2008). Figure S.3 depicts three such features — node degree (by node family), the number of partners shared by each dyad, and the minimum geodesic distance between nodes in the network. In each panel, the red line traces the true distribution of these network statistics, while the black vertical bars track their distribution across 100 network replicates, each generated using the estimated coefficients. If the latter are correctly estimated, network replicates should have characteristics that reflect that on which the estimation is based, and the red line should fall squarely within each vertical black rectangle. Overall, network characteristics are well recovered by our model, although recovery of the degree distribution for the bill nodes (i.e., the largest family) is less accurate than that of the legislator nodes (i.e., family with fewer vertices).

To evaluate scalability of our approach, we conduct simulations under the "medium" difficulty scenario, as described above. We hold all conditions constant, and increase the total size of the vertex set from 300 to 15,000, keeping twice as many vertices in the largest family as in the smallest family. In all instances, we let the models run until convergence, using the stochastic variational inference procedure described above (in

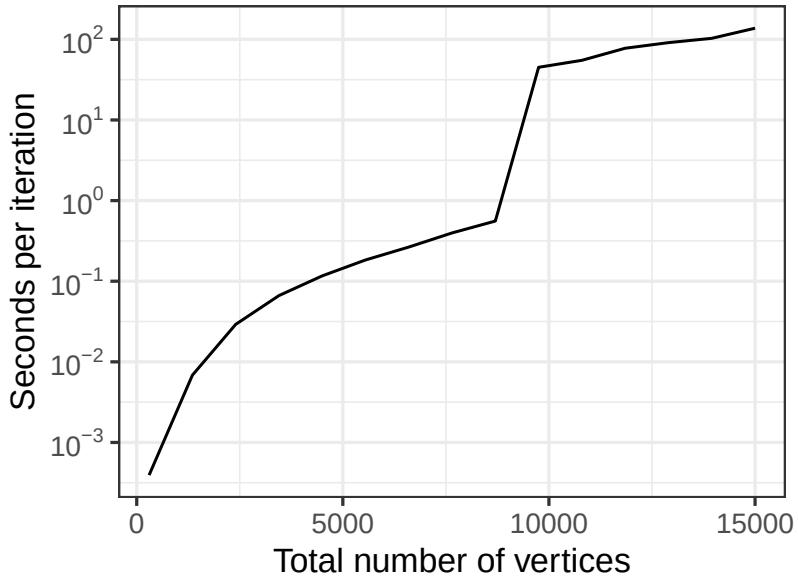


Figure S.4: **Time per iteration for networks of different sizes**: For medium difficulty scenario networks, the plot shows time per iteration (in seconds) taken to fit our model to networks of different sizes.

each iteration, we sample 40% of nodes). Models took between 5 seconds and 9 hours to fully estimate, taking anywhere between 100 and 450 iterations to converge. As the time to convergence is affected by the stochastic nature of the estimation, Figure S.4 presents the time per iteration (in seconds) taken to fit our model to networks of different sizes. Overall, although time per iteration increases as the network size grows, even the largest network in our simulation can be reliably fit in under 10 hours on a desktop computer.

Finally, we evaluate the frequentist properties of our estimates of uncertainty in regression parameters by evaluating the extent to which they reflect the variability we can expect from repeated network sampling. To do so, we sample 100 networks from each of our 6 scenarios, for a total of 600 simulated networks. Figure S.5 shows, for each simulation scenario, the difference between our standard errors and the standard deviation across coefficients estimated on each of the network replicates. For small networks, our standard errors can be conservative — particularly for the set of coefficients associated with the smaller group of Senator nodes. As the number of nodes increases, however, our estimated uncertainty more accurately reflects the variability we would expect to see under repeated sampling.

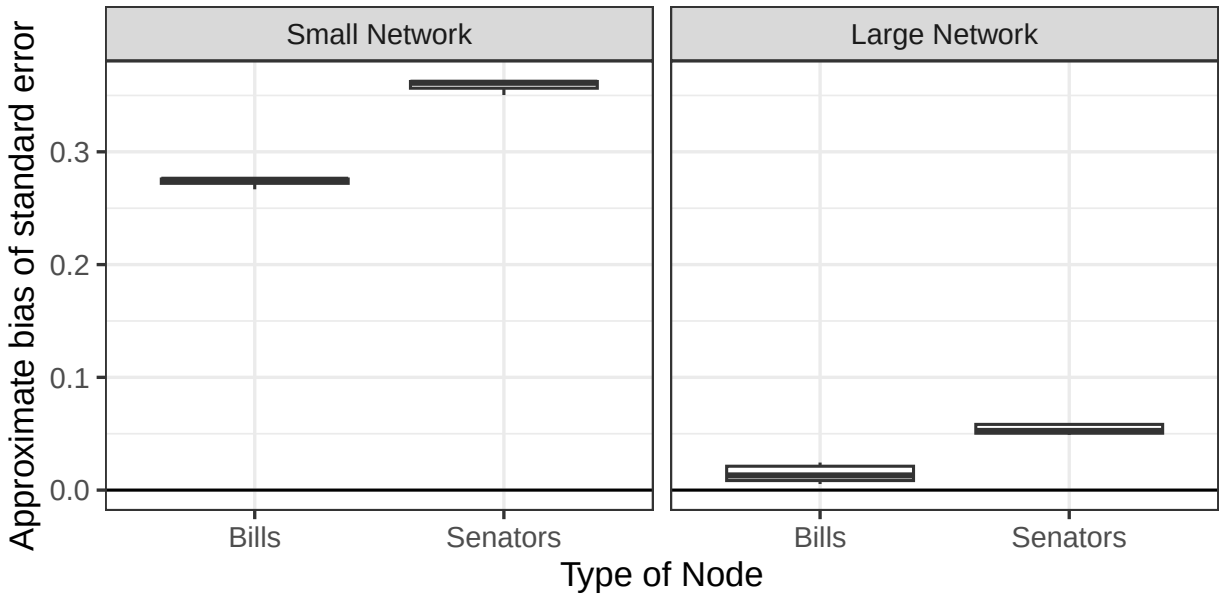


Figure S.5: **Approximate Bias in Standard Errors**: For each simulated network, the figure shows the extent to which our approximate standard errors differ from the standard deviation of coefficients estimated on simulated networks.

S.3 Additional empirical results

S.3.1 Cosponsorship degree distributions 107th Senate

Figure S.6 presents Senator and Bill degree distributions from the 107th Congress. Bipartite degree distribution calculations differ from unipartite ones in that they are separately conducted for each family, so that senators can display different degree distributions compared to bills. Previous studies of cosponsorship patterns in the U.S. congress have found this to be the case (e.g. Fowler, 2006), and our data reveal similar differences.

Figure S.6 displays a summary of these distributions, plotting the midpoints of the degree histograms for each vertex type. We plot both degrees and their observed relative frequencies (expressed as percentages) in the log scale. When degree distributions follow the common power-law distribution (whereby $p(x) \propto x^{-\lambda}$ for a given degree x and $\lambda > 0$) that many other networks exhibit, such log-log plots tend to align with negatively-sloped linear predictions.

This is clearly the case for the degree distribution of bills, depicted on the left panel of Figure S.6. The

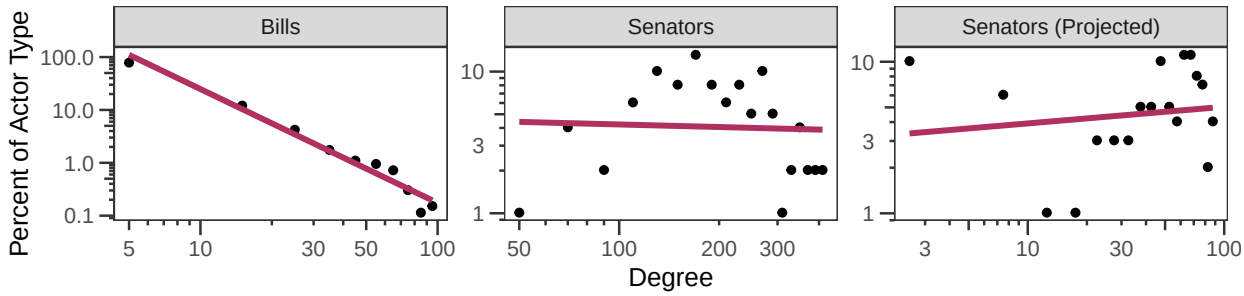


Figure S.6: **Bill and Senator degree distributions.** Degree distributions presented for Bills (left) and Senators (center, in the bipartite network, and right, in the projected network) with a power law distribution overlaid as a red curve (approximated by a linear log-log model). Whereas the power law fits the bill degree distribution quite well, the degree distributions among senators (both in the bipartite graph and in the projected network) differ dramatically from it, illustrating the kind of heterogeneity that can be lost when aggregating over bills in the projection from bipartite to unipartite.

plot also includes a red line with the predictions of linear log-log model, which shows a good approximation to the pure power-law model if the estimated slope is negative. For bills, the fit of a power-law is almost perfect, indicating that while most bills tend to attract few cosponsors, there is a long and heavy tail of bills attracting a large number of them. It is precisely this heterogeneity that can result in substantial aggregation bias when projecting the originally bipartite network. It also means that the cosponsorship network is likely to exhibit *scale invariance* for the distribution of bill cosponsors (i.e., we can expect to see a similarly shaped degree distribution if we consider a subset of bills) — justifying the analysis of subnetworks formed by sampling the set of bills (as we do below).[1]

In contrast, the degree distributions of Senators (i.e., the distributions over the number of bills Senators cosponsor) are far from being accurately described by a power-law. This is indicated by the fact that the red line does not fit to the points well on the central and right panels of Figure S.6. Indeed, these distributions are quite different from that of bills, suggesting that there is substantial heterogeneity in the number and strength of connections between senators. This highlights the importance of considering the entire set of senators when studying the structural characteristics of the cosponsorship network, and considering any subset of legislators could result in a misrepresentation of the collaboration network. This difference with

[1] As pointed out by Fowler (2006), the distribution also exhibits an interesting deviation from a typical power-law right around the 50 senator mark, indicating the strategic value of having a majority of senators cosponsoring a piece of legislation.

respect to the degree distribution of bills also highlights the kind of information that is lost when bills are aggregated over in the process of projecting from bipartite to unipartite networks, as is also evident on the right-most panel of Figure S.6, which appears now to be a mixture of two distributions.

The bipartite network also can accomodate certain types of statistics for descriptive analysis that is not applicable for unipartite networks. They include within-family edge-shared partners and family-specific k-stars. Conversely, several common network statistics for unipartite graphs do not apply in bipartite network settings, such as triangles (bipartite graphs cannot have triangles (Prömel, Schickinger, and Steger, 2002)), or must be adjusted, such as path lengths (which must be even) or minimum degrees (Liu and Ma, 2018). These considerations will play a role when conducting posterior predictive checks of model fit, as they typically rely on evaluating how well a model captures these and other structural features of the modeled network.

S.3.2 Model performance comparison

We compare our proposed approach to the most popular, readily available alternative model for bipartite networks: the bipartite ERGM, implemented in the R package `ergm` (Handcock et al., 2023; Hunter, Handcock, et al., 2008; Krivitsky et al., 2023). The bipartite ERGM uses a set of constraints and bipartite-specific network statistics to adapt the canonical, one-model model to bipartite networks. Our goal is to evaluate how well the bipartite ERGM can predict the cosponsorship network of the 107th Senate, using network statistics only. We then compare it to how biMMSBM can fit the same data using only the blockmodel and latent mixed membership vectors.

While we tried fitting the ERGM to the full dataset, we found all of the specifications we tried resulted in failed convergence; the lack of scalability appears to be a major limitation of ERGM. As a result, we focus on a subgraph formed by a random sample of 10% of observed edges. To this subgraph, we fit a model that includes a term for the edge density, the census of 3-stars among senators (i.e., stars involving exactly 1 bill and three senators), a geometrically weighted census of dyad-shared partners among bills (i.e., the distribution over numbers of shared senators for any pair of bills), and geometrically weighted degree

distributions for both senators and bills.[2] The latter terms are so-called dyad-dependent terms, and are defined specifically for bipartite networks (Wang, Sharpe, et al., 2009; Wang, Pattison, and Robins, 2013). After fitting the ERGM, all measures of MCMC performance indicated convergence.

For biMMSBM, we use 5 latent communities in each family (a number arrived at by evaluating the AUROC generated by alternative models, as we do in the main analysis. We compare models with $2-2$, $3-3$, $3-5$, and $5-5$ latent communities).

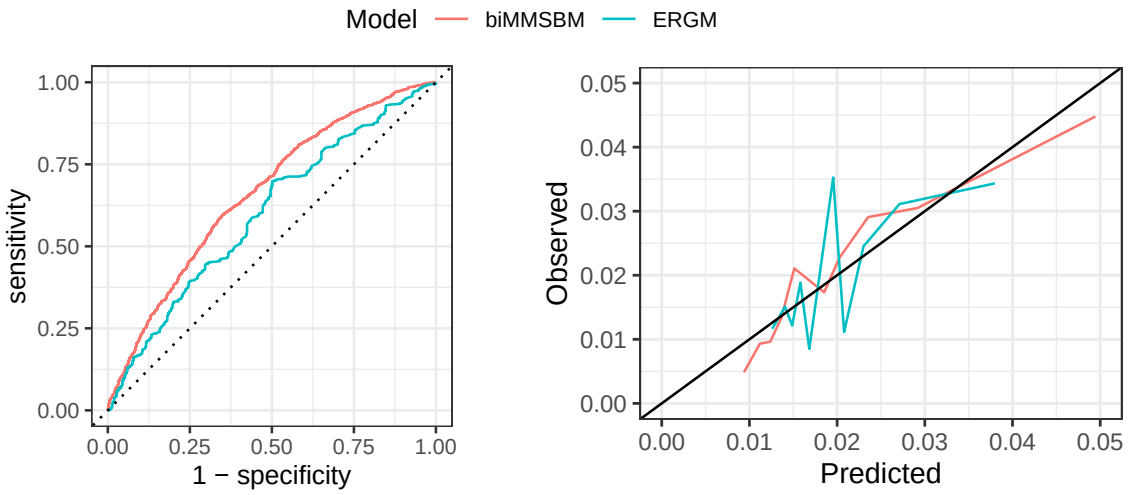


Figure S.7: Measures of predictive accuracy of bipartite ERGM and biMMSBM on subset of cosponsorship network in the 107th Senate. The left panel shows the ROC curves for the ERGM (blue) and biMMSBM (salmon); curves further from the 45-degree reference line indicate better model classification accuracy. The right panel shows the calibration of the same models, with lines closer to the 45-degree line indicating a better match between predicted edge probabilities and observed edge proportions (that is, better calibration). Using both criteria, biMMSBM offers a better predictive fit to the cosponsorship data.

Figure S.7 offers evidence of better predictive accuracy obtained by biMMSBM on this network, as indicated by a both a higher overall AUROC (0.66 vs. 0.59 obtained by the ERGM; left panel), and far better calibration of predicted probabilities (Platt et al., 1999, after running a Platt correction for both sets of predictions, []), as indicated by the alignment of the biMMSBM set of predicted probabilities with the

[2]We arrived at this particular specification through *much* trial and error, iterating over specifications that invariably hung or failed to converge. This failure-prone process, often elided from descriptions of empirical exercises that rely on ERGM-type models, is sometimes touted as a feature. Arriving at a specification that works, however, can only offer a weak proof of existence, and even a perfectly specified model can result in an ill-defined probabilistic models (see, for instance, Chatterjee and Diaconis, 2013).

corresponding empirical proportions of edges (right panel). These are both in-sample measures of fit, as the fragility of the ERGM estimation prevented us from performing an out-of-sample evaluation.

S.3.3 Goodness of Fit

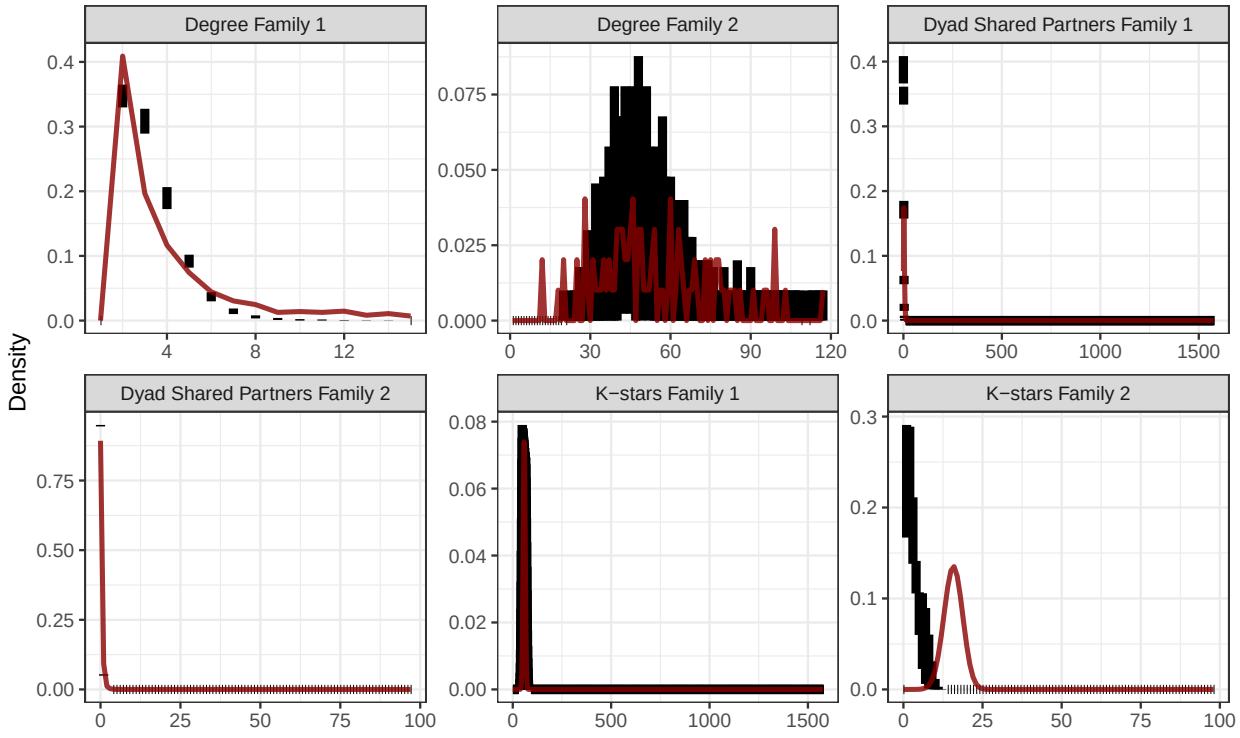


Figure S.8: **Posterior predictive goodness-of-fit checks, out-of-sample**. Vertical black rectangles represent the interquantile range across 50 replicate networks. The red line in each panel denotes the observed value in a network formed by a random 25% sample of cosponsorship decisions during the 107th Senate. The model generally replicates structural features well, shown by overlap between black bars and red lines. However, k-stars of bills are consistently underpredicted in the out-of-sample set.

We assess model performance by generating 50 replicates of the out-of-sample network from our model's posterior predictive distribution. We compare network-level statistics (e.g., degree distributions of senators and bills, family-specific distributions of edge-shared partners and k-stars) between these replicates and our test set. Figure S.8 displays the results, with black bars showing the interquartile range across replicates and red lines indicating observed values in the test set. The model generally fits the out-of-sample

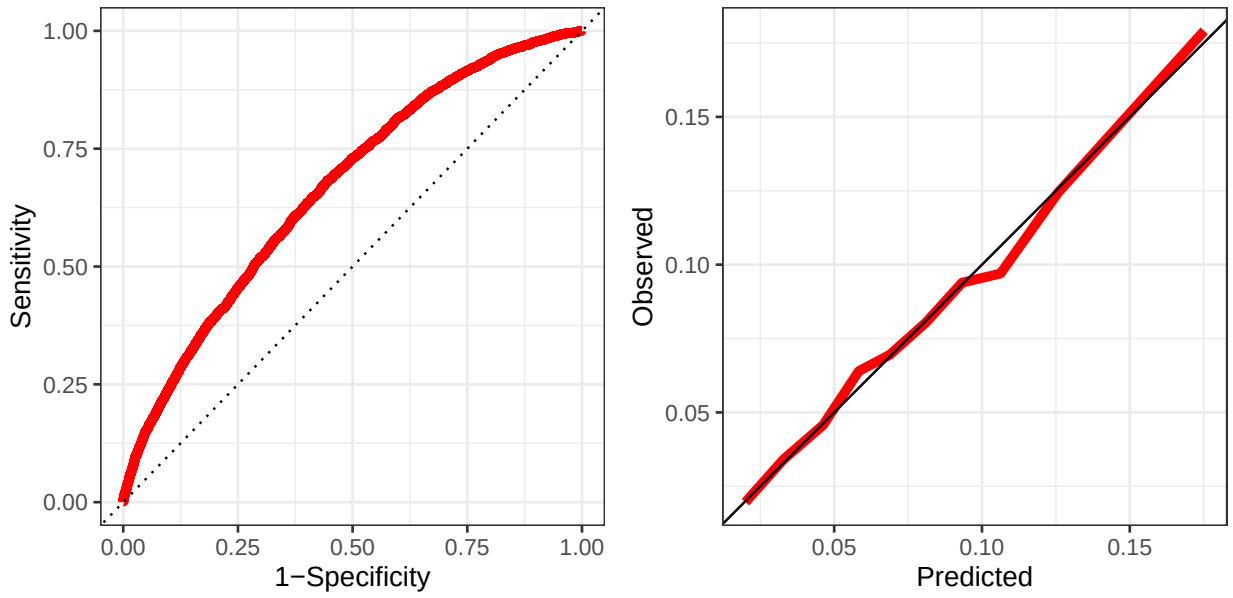


Figure S.9: **Out-of-Sample goodness-of-fit based on edge prediction quality:** The figure demonstrates that our biMMSBM model with $K_1 = K_2 = 3$ latent groups per family can adequately predict observed edges in the bipartite cosponsorship network. The left panel shows the Receiver-Operating Characteristic Curve in red, with curves closer to the upper-left corner indicating better predictive accuracy (here, the area under the curve is 0.70, out of a maximum of 1). The right panel shows the calibration of predicted probabilities (after standard Platt correction) in red, aligning with the 45-degree line for better predictive calibration.

network well, with most red lines falling within their corresponding black bars, capturing structural features without explicit specification. However, it tends to underpredict the distribution of k-stars among bills. The model also demonstrates high predictive accuracy and calibration, predicting observed edges with higher probabilities 70% of the time, as shown in Figure S.9.

Additionally, besides overall network statistics, we assess predictive quality using two metrics: accuracy, quantified by the area under the receiver operating characteristic curve (ROC), and calibration, comparing observed frequencies in a test set with model-predicted probabilities.

In Figure S.9, we show the out-of-sample ROC (left panel) as well as the prediction probability calibration plot (right panel).[3] The area under the ROC curve is .7 (with the red curve on the left bending away

[3]We evaluate calibration after applying the standard Platt correction, which fits a logistic regression of observed outcomes on the uncorrected predicted probabilities, and use the transformed scores (Platt et al., 1999; Rosenman, McCartan, and Olivella, 2023).

from the dashed diagonal), and the model's predictions are well calibrated (with the red curve on the right almost aligning with the solid black diagonal).

S.3.4 Model outputs

S.3.4.1 Group memberships

Senior Democrats	**Senior Republicans**	**Junior Power Brokers**
Kennedy, Edward M. [MA]	Helms, Jesse [NC]	Corzine, Jon [NJ]
Sarbanes, Paul S. [MD]	Nickles, Don [OK]	Carnahan, Jean [MO]
Inouye, Daniel K. [HI]	Craig, Larry E. [ID]	Clinton, Hillary Rodham [NY]
Byrd, Robert C. [WV]	Hatch, Orrin G. [UT]	Dayton, Mark [MN]
Akaka, Daniel K. [HI]	Lott, Trent [MS]	Carper, Thomas R. [DE]
Biden Jr., Joseph R. [DE]	Thurmond, Strom [SC]	Stabenow, Debbie [MI]
Leahy, Patrick J. [VT]	Gramm, Phil [TX]	Miller, Zell [GA]
Hollings, Ernest F. [SC]	Grassley, Charles E. [IA]	Bayh, Evan [IN]
Dodd, Christopher J. [CT]	Roberts, Pat [KS]	Chafee, Lincoln D. [RI]
Baucus, Max [MT]	Cochran, Thad [MS]	Feingold, Russell D. [WI]

Table S.2: **Senators with largest mixed-membership probabilities in each latent group.**

Uncontroversial	Bipartisan Resolutions	Contentious
SE_107_169 Resolution relative to the death of the Honorable Mike Mansfield, formerly a Senator from the State of Montana.	SJ_107_1 Joint resolution proposing an amendment to the Constitution of U.S. relating to voluntary school prayer.	SN_107_1548 Bill to allow Director of CDC to award a grant to create/maintain a website with bioterrorism information.
SJ_107_22 Joint resolution expressing the sense of Senate and House regarding the terrorist attacks launched against the Unites States on September 11, 2001.	SE_107_82 Resolution authorizing production of records by Permanent Subcommittee on Investigations of Committee on Governmental Affairs.	SN_107_2842 Bill to amend Older Americans Act of 1965; authorize appropriations for demonstration projects to provide supportive services to older individuals in NORCs.
SE_107_292 Resolution expressing support for Pledge of Allegiance.	SE_107_54 Resolution authorizing expenditures by the committees of the Senate.	SN_107_2899 Atchafalaya National Heritage Area Act
SE_107_354 Resolution relative to the death of Paul Wellstone, a Senator from the State of Minnesota.	SE_107_10 Resolution notifying the House of Representatives of the election of a President pro tempore of the Senate.	SN_107_2918 Bill designate USPS at 380 Main Street in Farmingdale, New York, as ”Peter J. Ganci, Jr. Post Office Building”.
SE_107_160 Resolution designating month of October 2001, as ”Family History Month”.	SE_107_77 Resolution to authorize production of records by Permanent Subcommittee on Investigations of Committee on Governmental affairs.	SN_107_1892 Bill to designate facility of USPS at 375 Carlls Path in Deer Park, New York, as ”Raymond M. Downey Post Office Building”.
SE_107_66 Resolution expressing sense of Senate regarding release of 24 US military personnel being detained by China.	SE_107_9 Resolution notifying POTUS of election of a President pro tempore.	SN_107_1721 Bill to designate building located at 1 Federal Plaza as ”James L. Watson United States Courthouse”.
SN_107_321 Dylan Lee James Act (Family Opportunity Act of 2004)	SE_107_28 Resolution to authorize testimony and legal representation in State of Idaho v. Fredrick Leroy Leas, Sr.	SN_107_1801 Bill to amend chapter 36 of title 39, United States Code, to provide for a permanent postal rate.
SN_107_677 Bill to amend Internal Revenue Code of 1986.	SE_107_84 Resolution to authorize representation by Senate Legal Counsel in Timothy A. Holt v. Phil Gramm.	SN_107_3176 Bill to amend the Internal Revenue Code of 1986.
SN_107_697 Bill to modernize financing of railroad retirement system and provide enhanced benefits to employees and beneficiaries.	SC_107_27 Concurrent resolution expressing the sense of Congress: 2008 Olympic Games should not be held in Beijing unless China releases all political prisoners [...].	SN_107_2634 Bill to establish within the National Park Service the 225th Anniversary of American Revolution Commemorative Program.
SN_107_1707 Bill to amend title XVIII of Social Security Act to specify update for payments under medicare physician fee schedule for 2002 and to direct the Medicare Payment Advisory Commission to conduct a study.	SJ_107_5 Joint resolution appointment of Walter E. Massey as a citizen regent of Board of Regents of Smithsonian Institution.	SN_107_3045 Bill to amend Federal Water Pollution Control Act to provide for protection and enhancement of environmental integrity and social and economic benefits of Finger Lakes Region in New York.

Table S.3: **Bills with largest mixed-membership probabilities in each latent group.**

S.3.4.2 Model estimated coefficients

Blockmodel estimates	1 Uncontroversial Bills	2 Bipartisan Resolutions	3 Contentious Bills
1 Democrat	1.0000	0.5969	0.4110
2 Republican	0.9998	0.4345	0.3084
3 Junior Power Brokers	1.0000	0.6187	0.4687

		Coefficient Name	Estimate	SE
Dyadic predictors				
		No reciprocity history	-5.8241	0.0596
		Log proportional reciprocity	1.7849	0.0101
		Shared committee	1.3246	0.0343
Model Summary Statistics				
Lower bound	-462.1155			
Number of dyads	260667			
% Obs. in Each Family 1 Block	0.223	0.374	0.403	
% Obs. in Each Family 2 Block	0.062	0.434	0.504	

Table S.4: **biMMSBM Estimated Coefficients: Blockmodel, Dyadic.** Point estimates and approximate standard errors of coefficients in the dyadic regression equation show that reciprocity norms and shared committee duties between bill sponsors and potential co-sponsors enhance co-sponsorship likelihood.

Group	Coefficient Name	Estimate	SE
1 Democrat			
1 Democrat	Intercept	6.0405	1.2906
	Seniority	0.6182	0.6851
	Ideology 1	-5.2906	1.3044
	Ideology 2	2.4779	1.2908
	Party-Republican	-3.3740	1.3014
	Sex-Male	0.6249	1.2853
2 Republican			
2 Republican	Intercept	5.7276	1.2883
	Seniority	0.4713	0.6851
	Ideology 1	7.1225	1.2959
	Ideology 2	1.5571	1.2907
	Party-Republican	4.2359	1.2910
	Sex-Male	1.1646	1.2847
3 Junior Power Brokers			
3 Junior Power Brokers	Intercept	22.2967	1.2868
	Seniority	-0.8769	0.6851
	Ideology 1	-3.9223	1.2951
	Ideology 2	-6.1953	1.2901
	Party-Republican	-1.6291	1.2908
	Sex-Male	-1.0508	1.2844

Table S.5: **biMMSBM Estimated Coefficients: Senator monadic.** Point estimates of coefficients in the scale of linear predictor, along with their corresponding approximate standard errors.

Group	Coefficient Name	Estimate	SE
1 Uncontroversial Bills			
1 Uncontroversial Bills	Intercept	1.7158	0.1418
	Topic:Legal	0.7253	0.1207
	Topic:Social programs Public goods	-0.0511	0.0733
	Topic:Security	0.9412	0.2094
	Topic:Gov operations	0.5139	0.1506
	Topic:Other	-0.8196	0.0814
	Sponsor Seniority	-0.0581	0.0065
	Sponsor Ideology 1	0.3522	0.2702
	Sponsor Ideology 2	0.0254	0.1571
	Sponsor Party-Republican	-0.2950	0.1996
	Sponsor Sex-Male	-0.4922	0.0971
	Second Phase	0.3407	0.0934
	Third Phase	0.0213	0.0625
2 Bipartisan Resolutions			
2 Bipartisan Resolutions	Intercept	3.8623	0.1251
	Topic:Legal	0.6426	0.1037
	Topic:Social programs Public goods	-0.1534	0.0654
	Topic:Security	0.9192	0.1632
	Topic:Gov operations	0.5713	0.1218
	Topic:Other	-1.3325	0.0728
	Sponsor Seniority	-0.0540	0.0057
	Sponsor Ideology 1	0.4956	0.2449
	Sponsor Ideology 2	-0.1777	0.1457
	Sponsor Party-Republican	-0.4312	0.1797
	Sponsor Sex-Male	-0.4587	0.0864
	Second Phase	0.6710	0.0803
	Third Phase	0.4733	0.0556
3 Contentious Bills			
3 Contentious Bills	Intercept	4.1264	0.1251
	Topic:Legal	0.6044	0.1037
	Topic:Social programs Public goods	-0.0325	0.0653
	Topic:Security	0.8943	0.1634
	Topic:Gov operations	0.6980	0.1219
	Topic:Other	-1.4841	0.0728
	Sponsor Seniority	-0.0672	0.0057
	Sponsor Ideology 1	0.6429	0.2447
	Sponsor Ideology 2	-0.2636	0.1454
	Sponsor Party-Republican	-0.6676	0.1795
	Sponsor Sex-Male	-0.5427	0.0864
	Second Phase	0.8368	0.0803
	Third Phase	0.7067	0.0555

Table S.6: **biMMSBM Estimated Coefficients: Bill monadic.** Point estimates of coefficients in the scale of linear predictor, along with their corresponding approximate standard errors.

S.3.5 Degree centrality and senator memberships

	Estimate	Std. Error	t value
Baseline: Power Brokers	6503.731	1026.778	6.3341
Senior Democrats	-1818.510	1761.168	-1.0326
Senior Republicans	-1406.040	1511.945	-0.9300

Note:

Multiple R^2: 0.01356 Adjusted R^2: -0.00678 F stat: 0.667

Table S.7: **Regression of senator between centrality on group assignment probabilities.** Baseline is Group 3, which is positively correlated with between centrality.

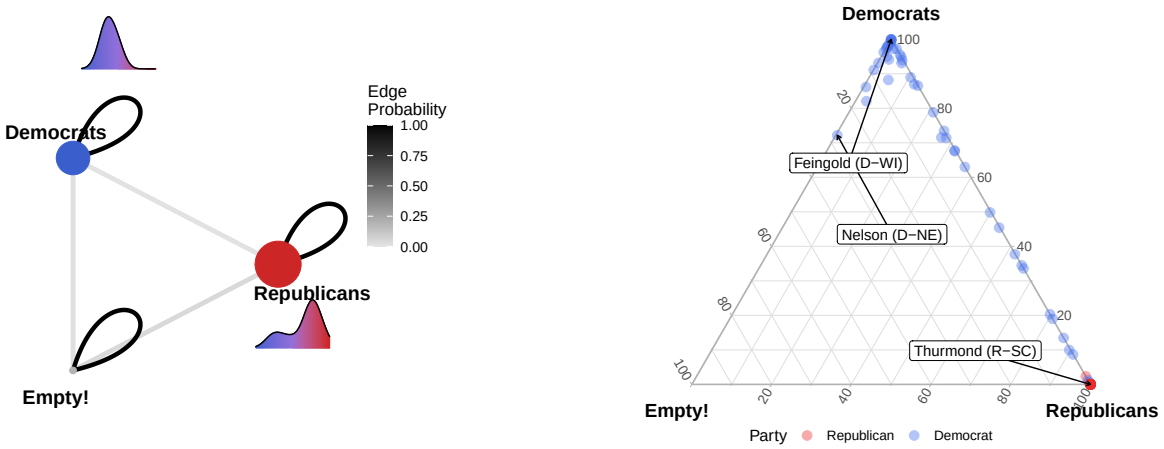
S.3.6 Comparison with the unipartite network model

We compare the results of our model against those of a unipartite network model — the modal current approach to studying bipartite data as discussed in our introduction. For direct comparison, we use a unipartite (and static) version of our model, known as `dynMMSBM` (Olivella, Pratt, and Imai, 2022). We project the bipartite network data onto a unipartite weighted network, in which the weight of edges between senators is given by the number of bills they cosponsor together.[4]

For fair comparison, we keep the model specification as similar to the one used in our bipartite network analysis as possible: three latent groups of senators, and the same set of senator-specific covariates. In addition, we use the estimated senator mixed-membership vectors from `biMMSBM` as the initial values of the corresponding vectors in the unipartite model. This increases our confidence that any difference in the learned grouping of legislators is a result of the data aggregation process, rather than the differences of model specification and estimation.

Figure S.10 presents the main quantities of interest from this comparison: the estimated blockmodel

[4] `dynMMSBM` can accommodate such weighted networks by using a binomial likelihood, where the weights are modeled as "successes" among a number of trials equal to the number of bills two senators could have cosponsored together.



(a) Unipartite Senator Groups Blockmodel

(b) Unipartite Senator Mixed-Memberships

Figure S.10: **Results from fitting a `MMSBM` to the projected cosponsorship network.** After fitting a unipartite `MMSBM` to the projected cosponsorship network in the 107th Senate, we recover a blockmodel of group interaction probabilities (left panel) and a set of group membership probabilities for Senators (right panel). As expected, the model produces groups that are mostly aligned with partisanship and ideology, as indicated by the ideological distributions of likely members depicted next to each estimated block. We also find evidence of high intra-group collaboration probabilities (especially among Democrats), and very low probabilities of connecting across partisan groups, painting a picture of a highly polarized session.

between the three discovered groups of senators in the left panel (i.e., the propensity of members of any of these groups to collaborate with a member of another latent group) and the composition of these groups in the right panel (i.e., the extent to which different senators are likely to be in any of the three groups).

In contrast to the results shown in Figures S.10a and S.10b offer a picture of a polarized, relatively non-cooperative session of Congress, divided clearly along partisan lines — exactly as we expected from our discussion of the issues brought about by the aggregation process involved in projecting a bipartite network onto a unipartite one (see Section 2.1).

While the unipartite model could accommodate three different groups of senators under its specification, the data support membership primarily into two of those latent blocks, leaving the third group essentially

empty. Furthermore, membership into these two blocks is strongly aligned with partisanship, as the bulk of Democratic senators (depicted as blue circles on the right panel of Figure S.10) concentrate on one vertex of the ternary plot (with liberal Democrat Barbara Boxer being estimated as the senator most likely to instantiate this latent group), while the majority of Republican senators (depicted as red circles) tend to concentrate on another. Those senators who tend to have a minimal likelihood of instantiating the "Empty!" block tend to be Republican ideologues, led by the notoriously conservative Jessee Helms (R-NC).

Moreover, the unipartite model identifies only a moderate amount of collaboration across the aisle, yielding the estimated probability of about .18 that a member of the latent "Democratic" group is connected in the projected unipartite network to a menber of the "Republican" group. In contrast, the probability of a connection between two members of the *same* partisan latent group is estimated to be about .99, as indicated by the darker-shaded loops in the blockmodel on the left of Figure S.10. While Democrats enjoyed a slight majority after senator Jeffords decided to leave the Republican party (suggesting that the cohesive majority managed to move legislation along), this picture of polarization painted by the unipartite model gives few clues as to why the 107th Senate was able to remain as productive as it was.

In sum, relying on the projected network of cosponsorships not only misses the rich and nuanced information about legislative collaboration that individual bills have to offer, but it also distorts what we can learn about the nature of legislative coalitions. Using a unipartite model to study naturally bipartite data results in an artificially inflated sense of clustering and group cohesion — a risk that becomes even more pressing when we are interested in group dynamics and polarization that drive processes of network formation.

S.3.7 Alternative model specifications

Separating block membership and covariate estimation Our proposed model allows for joint inference of block membership and predictor effects. It also allows for inclusion of different predictors. To illustrate these possibilities, we restrict the `biMMSBM` model to infer only block membership or only predictor effects. Findings are briefly summarized below:

1. **Block membership only**. Here no predictors contribute to variation in cosponsorship; instead information is highly concentrated on the block memberships, which are the same for every senator/bill for

each respective family of blocks. We find that senators who fall into any of the senator latent groups are likely to have a tie with a bill that instantiates the second bill latent group (ranging 0.613-0.765); all other block-to-block interactions are limited (ranging 0.004-0.042). This model suggests low levels of tie formation for bills that fall into any latent group outside of group 2 — a departure from the original joint estimation of results in our main model. The full table of model estimates is shown in Table S.8.

2. **Same-state effects**. In this setting, we incorporate a dyadic predictor for same-state membership between a senator a sponsor of a bill. In this specification, the same-state indicator appears as a strong predictor of a cosponsorship link. Despite this, it is still possible to discern similar results with respect to the discovered groups of legislators and bills. The model estimates are shown in Tables S.9-S.12.

Coefficient	Standard.Error
Monadic: Senator intercepts	
10.6352	0.4101
6.5183	0.4100
6.0022	0.4097
Monadic: Bill intercepts	
2.2122	0.0233
0.0932	0.0235
1.5949	0.0239

	Blockmodel		
	Bill 1	Bill 2	Bill 3
Senator 1	0.0103	0.7649	0.0020
Senator 2	0.0122	0.7555	0.0180
Senator 3	0.0041	0.6130	0.0416

Table S.8: **Only Block Membership Full Model Coefficients.** Coefficient point estimates in linear predictor scale, with corresponding approximate standard errors.

Table S.9: Same-State Full Model Coefficients: Senator Monadic. Coefficient point estimates in linear predictor scale, with corresponding approximate standard errors.

Group	Coefficient.Name	Estimate	SE
Senator predictors			
1 Democrat	Intercept	9.6722	1.2719
	Seniority	0.0229	0.3889
	Ideology 1	-6.7897	1.2818
	Ideology 2	2.6573	1.2866
	Party-Republican	-4.8794	1.4720
	Sex-Male	-0.6568	1.2605
2 Junior Power Brokers	Intercept	20.1595	1.2715
	Seniority	-0.9228	0.3889
	Ideology 1	-1.0970	1.2781
	Ideology 2	-7.2593	1.2872
	Party-Republican	-1.3287	1.3154
	Sex-Male	-1.1748	1.2605
3 Republican	Intercept	4.2601	1.2716
	Seniority	0.8223	0.3889
	Ideology 1	1.7553	1.2792
	Ideology 2	4.7461	1.2862
	Party-Republican	1.6111	1.3155
	Sex-Male	1.0517	1.2606

Table S.10: Same-State Full Model Coefficients: Bill Monadic. Coefficient point estimates in linear-predictor scale, with corresponding approximate standard errors.

	Group	Coefficient.Name	Estimate	SE
Bill predictors				
	1 Contentious Bills	Intercept	11.2669	0.1137
		Topic:Legal	0.8587	0.0303
		Topic:Social programs Public goods	0.8151	0.0555
		Topic:Security	0.8096	0.0709
		Topic:Gov operations	0.5672	0.0675
		Topic:Other	-5.5515	0.0981
		Sponsor Seniority	-0.1431	0.0147
		Sponsor Ideology 1	0.3879	0.5242
		Sponsor Ideology 2	-0.5291	0.3436
		Sponsor Party-Republican	-1.3865	0.3804
		Sponsor Sex-Male	-2.6192	0.1364
		Second Phase	3.6591	0.6810
		Third Phase	1.9508	0.1370
	2 Uncontroversial Bills	Intercept	5.5493	0.1362
		Topic:Legal	1.1332	0.0564
		Topic:Social programs Public goods	0.7925	0.0676
		Topic:Security	0.0947	0.1242
		Topic:Gov operations	0.0754	0.0963
		Topic:Other	-3.3246	0.0793

Table S.10: Same-State Full Model Coefficients: Bill Monadic. Coefficient point estimates in linear-predictor scale, with corresponding approximate standard errors. *(continued)*

CoefficientFamily	Group	Coefficient.Name	Estimate	SE
		Sponsor Seniority	-0.0827	0.0166
		Sponsor Ideology 1	-1.0612	0.5530
		Sponsor Ideology 2	0.0791	0.3645
		Sponsor Party-Republican	0.2602	0.4025
		Sponsor Sex-Male	-1.7541	0.1588
		Second Phase	1.5149	0.8123
		Third Phase	-0.0066	0.1424
	3 Bipartisan Resolutions	Intercept	10.5539	0.1137
		Topic:Legal	0.7711	0.0308
		Topic:Social programs Public goods	0.7775	0.0556
		Topic:Security	1.2102	0.0711
		Topic:Gov operations	0.9030	0.0678
		Topic:Other	-4.9057	0.0979
		Sponsor Seniority	-0.0878	0.0147
		Sponsor Ideology 1	0.5021	0.5239
		Sponsor Ideology 2	-0.0069	0.3433
		Sponsor Party-Republican	-1.6197	0.3802
		Sponsor Sex-Male	-2.6768	0.1365
		Second Phase	3.8117	0.6807
		Third Phase	0.9561	0.1368

Table S.11: Same-State Full Model Coefficients: Blockmodel.

	1 Contentious Bills	2 Uncontroversial Bills	3 Bipartisan Resolutions
1 Democrat	0.3652	0.9819	0.5175
2 Junior Power Brokers	0.1399	0.9939	0.2326
3 Republican	0.1088	0.9918	0.1960

Table S.12: Same-State Full Model Coefficients: Dyadic.

Coefficient.Name	Estimate	SE
Dyadic predictors		
No reciprocity history	-1.5580	0.0196
Log proportional reciprocity	0.4354	0.0044
Shared committee	0.6560	0.0208
Co-state	1.8788	0.0430