Supplementary Online Appendix
Endogenous Social Connections in Legislatures

Marco Battaglini

Cornell University and EIEF

battaglini@cornell.edu

Eleonora Patacchini

Cornell University

ep454@cornell.edu

Edoardo Rainone

Bank of Italy

edoardo.rainone@bancaditalia.it.

A.1 Extensions of the model omitted from the paper

A.1.1 Allowing for atomistic players

The analysis presented in the previous sections is based on the assumption that the players of the game described in Section 2.1 are "price takers," that is "non-atomistic" players who establish connections taking the effectiveness of the other players as given. We can, however, extend the analysis to study environments in which there are both non-atomistic players (the followers), and atomistic players (the leaders) who select social connections that affect the behavior of all other players, and who may be strategic about their choice of connections. The NCE introduced in Section 3.1 remains a key tool to study these environments: the idea is that we can apply the NCE to the followers to solve for their network connections *given* the connections of the large players. The use of the NCE allows us to drastically reduce the complexity of the problem by modeling the web of links among followers. We can focus on the much smaller network of links among the leaders, by either directly estimating it link by link or by modeling their interaction as a game. This is a general approach that significantly extends the applicability of the techniques presented above.

Assume that there are two types of legislators: common legislators $\mathcal{O} = \{1, ...q\}$ and leaders $\mathcal{M} = \{1, ..., m\}$. We define $\mathcal{N} = \mathcal{O} \cup \mathcal{M}$. Common legislators can be partitioned in m groups: $\{M_l\}_{l=1}^m$, each group M_l is associated to a leader l. The leader can be the leader of a political faction, or any other agent with an institutional or prominent role. The effectiveness of a player is a function of social connectedness and effort as in (1) in the paper. For a follower $i \in \mathcal{O}$, we define social connectedness as:

$$s_i = \sum\nolimits_{j \in \mathcal{O}} A_{i,j} \cdot g_{i,j} E_j(G, \varepsilon) + \sum\nolimits_{l \in \mathcal{M}} B_{i,l} \cdot g_{i,l} E_l(G, \varepsilon), \tag{A.1}$$

Equation (A.1) corresponds to equation (2) in the paper, except that now the importance of social links depends on the type of agent to whom i chooses to associate. For example, we may allow $A_{i,j} > A_{i,k}$ when j belongs to the same group as i, but k does not; similarly we may allow a follower to value a link with a leader more than a link to another follower. For leaders, we allow social connectedness s_l to depend on the type of legislator to whom l connects. Specifically, for any leader $l \in \mathcal{M}$, we can assume:

$$s_l = \kappa_G \sum\nolimits_{j \in M_l} g_{l,j} \cdot E_j(G, \varepsilon) + \kappa_{NG} \sum\nolimits_{j \notin M_l} g_{l,j} \cdot E_j(G, \varepsilon) + \kappa_L \sum\nolimits_{k \in \mathcal{M}} g_{l,k} E_k(G, \varepsilon). \tag{A.2}$$

so a leader values a link to followers of his/her own group differently than those to followers of other groups or to other leaders.

The game proceeds as follows. At $t = 2$, all legislators chose the respective levels of effort l_i for $i \in \mathcal{N}$, taking the entire network as given. At $t = 1$, the followers form their links according to the NCE: followers are forward looking, but take the effectiveness of all other lawmakers and the links formed by the leaders as given. At $t = 0$, the leaders select their links. The leaders are forward-looking, acting as Stackelberg "first movers" and internalize the effect of their own actions on the NCE. To formalize this model, it is useful to make the notation compact defining:

$$D_{i,j} = \begin{cases} A_{i,j} & i, j \in \mathcal{O} \\ B_{i,j} & i \in \mathcal{O},\, j \in \mathcal{M} \\ \kappa_G & i \in \mathcal{M},\, j \in M_i \\ \kappa_{NG} & i \in \mathcal{M},\, j \notin M_i \\ \kappa_L & i \in \mathcal{M},\, j \in \mathcal{M}. \end{cases}$$

With this, we can rewrite (A.1) and (A.2) as:

$$s_i = \sum_{j \in \mathcal{N}} D_{i,j} \cdot g_{i,j} \cdot E_j(G, \varepsilon).$$

Given the social network G, the equilibrium levels of effort are equal to:

$$E_i(G, \varepsilon) = \delta \cdot \left[\sum_{j \in \mathcal{N}} D_{i,j} \cdot g_{i,j} \cdot E_j(G, \varepsilon) \right] + \varepsilon_i. \tag{A.3}$$

Substituting the optimal effort in the legislators expected utility, we have:

$$U^i(G, \varepsilon) = \alpha\delta \left[\sum_{j \in \mathcal{N}} D_{i,j} \cdot g_{i,j} \cdot E_j(G, \varepsilon) \right] + \varepsilon_i.$$

Let $G(\mathbf{g}_{\mathcal{M}})$ be the network when the leaders select links $\mathbf{g}_{\mathcal{M}} = \{g_{l,k}\}_{l \in \mathcal{M},\, k \in \mathcal{N}}$. An interior choice for the ordinary legislators $i \in \mathcal{O}$ maximizes:

$$\alpha\delta \cdot \left[\sum_{j \in \mathcal{N}} D_{i,j} \cdot g_{i,j} \cdot E_j(G, \varepsilon) \right] - \frac{\lambda}{(1+\lambda)} \left(\frac{g_{i,j}}{\theta_{i,j}} \right)^{1+\frac{1}{\lambda}}.$$

So we have:

$$g_{i,j}(\mathbf{g}_{\mathcal{M}}) = (\theta_{i,j})^{1+\lambda} \left[\alpha\delta \cdot D_{i,j} \cdot E_j \right]^{\lambda}. \tag{A.4}$$

This defines $\mathbf{g}_{\mathcal{N}}(\mathbf{g}_{\mathcal{N}}) = (g_{i,j}(\mathbf{g}_{\mathcal{M}}))_{i \in \mathcal{O}, j \in \mathcal{N}}$. For $i \in \mathcal{O}$, let $M^{-1}(i)$ be i's leader (so that $i \in M_{M^{-1}(i)}$)

$$E_i = \alpha^{\lambda} (\delta)^{1+\lambda} \cdot \left[\sum_{j \in \mathcal{N}} D_{i,j} \cdot g_{i,j}(\mathbf{g}_{\mathcal{M}}) \cdot E_j \right] + \varepsilon_i.$$

For $l \in \mathcal{M}$:

$$E_l = \delta \cdot \left[\kappa_G \sum_{j \in M_l} g_{l,j} \cdot E_j + \kappa_{NG} \sum_{j \notin M_l} g_{l,j} \cdot E_j + \kappa_L \sum_{k \in \mathcal{M}} g_{l,k} E_l \right] + \varepsilon_i,$$

where $g_{l,k}$ is given by $\mathbf{g}_{\mathcal{M}}$. We solve for $E(\mathbf{g}_{\mathcal{M}})$ and we then find the network from $\mathbf{g}_{\mathcal{M}}$ and:

$$g_{i,j}(\mathbf{g}_{\mathcal{M}}) = (\theta_{i,j})^{1+\lambda} \left[\alpha\delta \cdot D_{i,j} \cdot E_j(\mathbf{g}_{\mathcal{M}})\right]^{\lambda} \text{ for } i \in \mathcal{O}. \tag{A.5}$$

To solve the game at stage 0, we need to specify the details of the leaders' strategic environments: whether they can select links simultaneously or sequentially; whether they can collude or coordinate, etc. We can now follow one of two approaches to solve for $\mathbf{g}_{\mathcal{M}}$, depending on the specifics of the environments in which the leaders interact. We have two possible approaches.

A.1.1.1 Direct estimation

The first approach is to make minimal assumption on how the leaders strategically interact. Given (A.5), we now derive the entire social network as a function of $\mathbf{g}_{\mathcal{M}}$: $G^*(\mathbf{g}_{\mathcal{M}}) = \left(g^*_{i,j}(\mathbf{g}_{\mathcal{M}})\right)_{i,j\in\mathcal{N}}$. Instead of specifying the details of how the leaders strategically interact, we can then leave $\mathbf{g}_{\mathcal{M}}$ as free variables and estimate them as parameters of the model. We can go back to the old algorithm, evaluating $z(E, \{\omega, \mathbf{g}_{\mathcal{M}}\})$, where now the vector of parameters to estimate is $\{\omega, \mathbf{g}_{\mathcal{M}}\}$, thus including $\mathbf{g}_{\mathcal{M}}$. We can define

$$z_i(\mathbf{E}, \{\omega, \mathbf{g}_{\mathcal{M}}\}) = E_i - \delta \cdot \left[\sum\nolimits_{j\in\mathcal{N}} D_{i,j} \cdot g^*_{i,j}(\mathbf{g}_{\mathcal{M}}) \cdot E_j\right] - \varepsilon_i,$$

and estimate the posterior distributions using Algorithm C defined in Section 4.2 of the paper. Such an estimation would be impossible with hundreds of players and dense networks, but may become feasible now because by using the NCE we can solve out for the social links of the followers.[81]

A.1.1.2 Modeling the leaders' behavior

The second approach is to specify a detailed game to describe how the leaders interact; and then solve for the entire game, thus obtaining predictions for the social connections of both the leaders and the followers. A convenient game form to model the leaders' interactions is to assume that they select their links sequentially, choosing their links in the order of their index $l = 1, ..., m$. Let $G(\mathbf{g}_{\mathcal{M}\setminus m})$ be the network whit the leaders up to the $(m-1)th$ be $\mathbf{g}_{\mathcal{M}\setminus m} = \{\mathbf{g}_l\}_{l=1}^{m-1}$. Now consider the mth leader. For simplicity, we assume here that the leader selects a vector of links to all other leaders $\mathbf{g_m} = (g_{m,1}, ..., g_{m,n})$, and a common link to all followers in his group $g_{m,G}$, and the other groups $g_{m,NG}$.[82]

[81] The direct estimation of the social network among the leaders is also possible if the number of leaders is large, but the network of social connections among them is sufficiently sparse (so that it is mostly constituted by links equal to zero). In these cases, machine learning techniques and rich datasets can be used to directly estimate the social networks among the leaders. Peng [2019], Battaglini et al. [2020a]), among others, present for a LASSO-based approach to estimate social networks in these cases. Here too the NCE is useful because, by solving out for the links among the followers, we can relax the constraint on how sparse the network among the leaders must be.

[82] The followers are anonymous for the leader, so it is natural to assume that s/he connects to them anonymously.

We can now solve for $\mathbf{g}_{\mathcal{M}}$ as a function of the other parameters of the model as follows. We assume that the links to each other leader is $\{0,1\}$ or, in other words, a leader either links with another leader or not; and the link to a group of followers M_l is also $\{0,1\}$, so a leader either links to all the followers of a given group M_l or not. The cost of forming a link to another leader is K_1; the cost of forming a link to a group of followers is K_2. Let $G(\mathbf{g}_{\mathcal{M}\backslash m})$ be the network when the leaders up to $m-1$ select $\mathbf{g}_{\mathcal{M}\backslash m} = \{\mathbf{g}_l\}_{l=1}^{m-1}$. Consider now leader number m. S/he solves the problem:

$$\max_{\mathbf{g}_m} \left\{ \begin{array}{c} \alpha\delta \cdot \left[\begin{array}{c} \kappa_G \sum_{j\in M_l} g_{m,j} \cdot E_j(G(\mathbf{g}_{\mathcal{M}}, \mathbf{g}_m), \varepsilon) \\ +\kappa_{NG} \sum_{j\notin M_l} g_{m,j} \cdot E_j(G(\mathbf{g}_{\mathcal{M}}, \mathbf{g}_m), \varepsilon) \\ +\kappa_L \sum_{k\in\mathcal{M}} g_{m,k} E_l(G(\mathbf{g}_{\mathcal{M}}, \mathbf{g}_m), \varepsilon) \end{array} \right] \\ -\sum_{k\in\mathcal{M}} K_1 \cdot 1_{m,k} - \sum_{k\in\mathcal{O}} K_2 \cdot 1_{m,k} \end{array} \right\}.$$

where $1_{m,k}$ is one if $g_{m,k} = 1$ and zero otherwise. This defines $\mathbf{g}_l^*(\mathbf{g}_{\mathcal{M}\backslash m})$. Proceed as above backward to define $\mathbf{g}_l^*(g_{\mathcal{M}\backslash l})$ for $l = 1, ...m$.
Once we have the equilibrium $G^* = \left(g_{i,j}^*\right)_{i,j\in\mathcal{N}}$ we can go back to the old algorithm, evaluating $z(E,\omega)$. We can define:

$$z_i(\mathbf{E}, \mathbf{g}_{\mathcal{M}}^*, \omega) = E_i - \delta \cdot \left[\sum_{j\in\mathcal{N}} D_{i,j} \cdot g_{i,j}^* \cdot E_j\right] - \varepsilon_i,$$

and estimate the posterior distributions using algorithm C defined in Section 4.2 of the paper.

Compared to the approach developed in the paper, the two approaches presented in this section allow for better differentiation of the roles played in the social network by different types of players, but they require more intrusive assumptions and they considerably complicate the analysis. Specifying *ex ante* the identity of the "leaders" and how they interact may be difficult to observe in practice. In the U.S., for instance, the speaker of the House may be the leader of his/her party, but this may depend on whether the same party has the majority in the House and/or the Senate; and whether that party also holds the Presidency. This approach is certainly even more challenging in other contexts: for example, when studying adolescents, or CEOs and corporate board members. Modeling atomistic and non-atomistic players, moreover, increases the computational complexity of the model. The analysis presented above shows that the "simple" approach in which all players are "price-takers" considerably improves the explanatory power of the model compared to models that ignore the endogeneity and unobservability of the social network. We leave for future research the investigation of whether allowing for atomistic non-price takers players improves the performance of the model even more.

A.1.2 Dynamic networks

In environments in which the agent's performance depends on social connections and we observe a measure of performance of the agents over long periods of time, it is natural to allow the social network to change over time. In these environments the social network at time t can be seen as a function of the network at time $t-1$. This may occur because it is cheaper to maintain a social connection than to form a new one, or because existing connections may make the formation of new connections easier (as when i knows j who know k, so it is easier for i to form a link with k). In these environments, moreover, forward looking agents would certainly anticipate the long term effect of connections.

While the model presented in the paper is static in the sense that the network is formed only once, some of the effects mentioned above are captured in the existing framework. As discussed in Section 2, the cost of forming a social link between i and j may depend on factors idiosyncratic to i and j through the term $h_{i,j}$: so if we know that i and j were previously socially connected, we can control for it when studying network formation at t. In the empirical application we use the alumni connection as a proxy for previously established connections, but depending on the environments we could have more information available. If we cannot observe proxies of social connections, we can still control for factors that may predict the existence of previous links, such as measures of demographic similarity or other variables. In our application, we control for the tenure of lawmakers because those that served in previous Congresses may have formed social connections among themselves. The model allows for the possibility of these effects; but it also allows the data to be used to assess if these variables are relevant in the formation of the social connections.

In addition, our model can be interpreted as the stage game of a more general dynamic model in which the network at time $t-1$ is taken as a state variable in the network formation at time t: in this more general model, the adjacency matrix $h_{i,j}^t$ used at time t is the network $g_{i,j}^{t-1}$ formed at $t-1$ (or more generally $h_{i,j}$ is a function of the network $g_{i,j}^{t-1}$). Given an initial observed adjacency matrix (when available), the model would endogenously account for it. Clearly this is a significantly more complex model than the one period version studied in this paper. Again, part of the complication lies in the fact that the formation of any link at t has externalities for all other links at t (as in this paper), but also now at $\tau \geq t$. We conjecture that in this dynamic environment the NCE can also play a key role in solving the model. We leave for future research the development of this important extension.

A.1.3 Negative spillovers

In the model presented above, i can only gain if j's effectiveness increases: if i and j are compatible (i.e. $\theta_{i,j} > 0$), then i can establish a link with j and benefit from j's effectiveness. If i and j are not compatible (say they have very different ideologies and they dislike each other), then i cannot establish a link with j, but j cannot hurt i.[83] There may be situations in which i does not want j's effectiveness to be high because j may actively use his effectiveness to contrast i. In this case, $g_{i,j} < 0$ independently from what i does. To allow for this possibility, we can introduce a variable $\varkappa_{i,j} = 1$ if i and j are enemies and zero otherwise. We can then modify the model assuming that if $\theta_{i,j} > 0$, then $\varkappa_{i,j} = 0$, so that if i can form a link with j, then j is not an enemy; but if $\theta_{i,j} = 0$, then $\varkappa_{i,j}$ can be 0 or 1. The link is now $(1 - \varkappa_{i,j})g_{i,j} - Z\varkappa_{i,j}$, so that if j is an enemy, then the effect of j's effectiveness on i is $-Z$. We can then estimate the parameters determining $\varkappa_{i,j}$ in the model as a function of the party affiliation and other homophily measures.

A.2 Additional proofs

Proof of the result in Example 2 of Section 3.3

First, consider an equilibrium with no connections. A necessary and sufficient condition for its existence is that a legislator, expecting no connections with the other players, finds it optimal to establish no connections as well. In this equilibrium, the effectiveness of an agent j is ε. Agent i finds it optimal not to link to $j = i + 1$ or $i - 1$ if $\alpha\delta\varepsilon - 1 \leq 0$, that is if $\varepsilon \leq 1/(\alpha\delta)$. Conversely, assume all legislators except i are fully connected. Then the equilibrium effectiveness of an agent j is $E = \frac{\varepsilon}{1-2\delta\overline{g}}$. Legislator i finds it optimal to connect to j if $\alpha\delta\frac{\varepsilon}{1-2\delta\overline{g}} - 1 \geq 0$, that is $\varepsilon \geq (1 - 2\delta\overline{g}) / (\alpha\delta)$. ■

A.3 Approximate Bayesian Computation

In this section, we detail the features of our ABC algorithm.

Prior Distributions. We adopt the following prior distributions for the parameters in model (17)-(19):

$$\begin{aligned}
\lambda &\sim U[0, \lambda_0], & (\psi, \gamma, \iota) &\sim N_{K_l+2}(\omega_0, \Omega_0), \\
\alpha &\sim U[0, 1], & \rho &\sim U[0, \varpi], \\
\eta_{i,r} &\sim N(0, \eta_0), & \sigma_\epsilon^2 &\sim TN_{\{0,\infty\}}(\sigma_0, \Sigma_0), \\
\beta &\sim N_K(\beta_0, B_0), & \zeta_r &\sim N(0, \sigma_\zeta), \\
& & \mu &\sim N(0, \mu_0),
\end{aligned}$$

where $U[\cdot]$, $TN_{\{a,b\}}(\cdot)$ and $N(\cdot)$ are the uniform, truncated normal (with a and b as lower and upper bounds), and normal distributions respectively. For our key parameters of interest measuring the social externality (ρ, λ and α), we adopt a uniform (uninformative) prior, as suggested in Smith and LeSage [2004] for spatial autoregressive models. Following Hsieh and Lee [2014], we adopt standard

[83] Indeed, i can benefit indirectly from j's effectiveness if there is a chain of connections such that, for example, j helps k who helps l who helps i.

normal priors for the parameters of covariates in the outcome and link formation equations (i.e. β, ψ, ι), and for $\eta_{i,r}$ and μ (if the model includes the unobservables as described in Section 6.3). The normal allows us to incorporate prior information regarding the variance-covariance matrix of the covariates' parameters in a natural way. We set the hyperparamethers as follows. For the parameters for which we have no prior information, we choose neutral values: λ_0 is set equal to five; η_0 (i.e., the variance of the prior for η) is set equal to one; σ_0 is set equal to zero; ϖ is set at 1; Σ_0 is a diagonal matrix with 0.1 on its diagonal elements; σ_ζ is set equal to 0.5. For other parameters we used available information to inform the prior as suggested by Kass and Wasserman [1996]: K is the number of controls in the outcome equation; K_l is the number of controls in the link formation equation; β_0 is set equal to the OLS point estimate obtained by regressing the controls on the outcome controlling for Congress fixed effects, and B_0 is set equal to the corresponding variance covariance matrix; ω_0 is set equal to the logit point estimates obtained by regressing the pairwise controls on the cosponsorship network entries, and Ω_0 is the correspondent variance covariance matrix. Hyperparameters in the prior distribution for the fixed effects ζs are given and fixed, differently from random effects (see Lancaster, [2004]; Rendon, [2013]).

Under the assumption that the social network G is observable and exogenous, conditions are generally imposed to guarantee an invertibility condition of G (see Kelejian and Prucha, [2010]), which in turn are sufficient for the existence of a unique equilibrium (see Calvo-Armengol et al. [2009]). The analogous condition in our theory of endogenous network formation is given by Proposition 2, stating a sufficient condition for the existence of a unique equilibrium. We therefore focus on a parameter space satisfying the condition of Proposition 2 that guarantees the existence of a unique equilibrium. To this goal, we extract values of λ, ρ, α, ψ, γ, ι and $\eta_{i,r}$ only if they satisfy Proposition 2, i.e. when $\delta < (1/\overline{\theta}) \cdot \left[1/\left((1+\lambda)\,\alpha^{\lambda}\overline{m}\right)\right]^{1/(1+\lambda)}$. Observe that ψ, γ, ι and $\eta_{i,r}$ are included in the formula because their values shape $\overline{\theta}$ and $\overline{m}$.

We should emphasize that the results are not sensitive to these assumptions about the prior distributions. The posterior distributions estimated in the empirical application are reported in Figures A.9-A.11.

Sampling Algorithm. The initial state of the Markov chain

$$\omega^{(1)} = [\lambda^{(1)}, \alpha^{(1)}, \eta^{(1)}, \beta^{(1)}, \psi^{(1)}, \gamma^{(1)}, \iota^{(1)}, \mu^{(1)}, \rho^{(1)}, \sigma_{\epsilon}^{(1)}, \zeta^{(1)}],$$

is set with all values equal to zero, except for $\beta^{(1)}$, $\psi^{(1)}$, and $\iota^{(1)}$. $\beta^{(1)}$ is set equal to the OLS point estimate obtained by regressing the controls on the outcome controlling for Congress fixed effects; $\psi^{(1)}$, and $\iota^{(1)}$ are set equal to the logit point estimates obtained by regressing the pairwise controls on the cosponsorship network entries.[84] To draw new values for each parameter (ω_i') at iteration t, we use a normal kernel, with mean equal to the current value and variance set at a

[84]The algorithm is robust to different starting values.

parameter-specific tuning parameter c:

$$N(\omega_{i,t}, c). \tag{A.6}$$

The decision rule for acceptance or rejection is described in Algorithm C (steps $C3$ and $C4$) in Section 4.2. Each step of the algorithm is run for each parameter, conditioning on the previous draws of the other parameters. Once every parameter has been updated, the algorithm moves to the next iteration.

To make the acceptance rate of the parameters' proposals as close as possible to 0.44 (which is optimal for one-dimensional proposals, see Roberts et al., [1997]; Roberts and Rosenthal, [2001]), we determine c with the following adaptive Metropolis-Within-Gibbs algorithm (see Roberts and Rosenthal, [2009]).[85] In the first phase, we allow c to change at each iteration t: c_t is decreased by a half percentage point if the algorithm presents an acceptance rate inferior to 20% in drawing new values; and is increased by half percentage point if the algorithm presents an acceptance rate superior to 80% in drawing new values. Namely:

$$\begin{array}{ll} \text{if}\, t_{A,i}/t \leq 0.2 & \text{then}\, c_{t+1} = c_t/1.005, \\ \text{if}\, t_{A,i}/t \geq 0.8 & \text{then}\, c_{t+1} = c_t \times 1.005, \\ \text{if } 0.2 \leq t_{A,i}/t \leq 0.8 & \text{then}\, c_{t+1} = c_t, \end{array} \tag{A.7}$$

where $t_{A,i}$ is the number of accepted draws at iteration t. The sequence c_t converges after the 10,000th iteration to a level c_∞. In the second phase, the parameter is set at its convergence level c_∞. This mechanism guarantees a bounded acceptance rate and convergence to optimal tuning. Figure A.12 reports the acceptance rate ($t_{A,i}/t$), which is the probability of moving from ω_i to ω_i', for each of our parameters over the $MCMC$ iterations. We observe that rates converge to values ranging from 40 to 85 percent, showing good mixing properties.

Our algorithm relies on the choice of the tolerance ν, the maximum acceptable distance between the simulated data from real data. Here too we proceed with a two-step procedure. In Step 1, we allow our algorithm to explore the tolerance space in the first 10,000 iterations. In Step 2, we then fix ν. Specifically, we use the following procedure.

First Step

C1 Start M parallel chains of length T with random initial vectors of parameters ω_m, with $m = 1, ..., M$, for each of them:

- C1.1 Propose to move from the current value ω to ω' according to a transition kernel $q(\omega \rightarrow \omega')$.
- C1.2 If $\varrho\left(z(\mathbf{E},\omega')\right) \leq \varrho\left(z(\mathbf{E},\omega)\right)$, proceed to the next step; else remain at ω; go to the first step.

[85] Our results are robust to the use of different adaptive algorithms, which are not reported for brevity.

- Calculate $h = \min\left(1, \frac{\pi(\omega')q(\omega' \to \omega)}{\pi(\omega)q(\omega \to \omega')}\right)$.

C1.3 Move to ω' with probability h, else remain at ω; go to the first step.

C2 Drop the first B observations of each chain.

C3 Compute the average distance $d_m = \frac{1}{T-B}\sum_{t=B+1}^{T} \varrho\left(z(\mathbf{E},\omega_m^t)\right)$, sort the chains according to d_m, and select the top τM chains.

C4 Set $\nu = prct^p \varrho\left(z(\mathbf{E},\omega_m^t)\right)$.

Second Step

C5 Set $\pi(.) = Pr(\omega|\varrho\left(z(\mathbf{E},\omega)\right) < \nu)$

C6 Start M parallel chains of length T with initial vectors of parameters ω_m drawn from π, with $m = 1, ..., M$, for each of them:

C6.1 Propose to move from the current value ω to ω' according to a transition kernel $q(\omega \to \omega')$.

C6.2 If $\varrho\left(z(\mathbf{E},\omega')\right) < \nu$, proceed to the next step; otherwise return to the first step.

C6.3 Calculate $h = \min\left(1, \frac{\pi(\omega')q(\omega' \to \omega)}{\pi(\omega)q(\omega \to \omega')}\right)$.

C6.4 Move to ω' with probability h, else remain at ω; go to the first step.

C7 Derive the posterior $Pr(\omega|\varrho\left(z(\mathbf{E},\omega)\right) < \nu)$.

ω_m^t is the value of ω at iteration t in the chain m, $prct^p$ is the p percentile function. In our benchmark estimation procedure, we set $M = 16$, $\tau = 0.75$, $p = 20$, $B = 1/4T$.

In this way, the algorithm moves in the first step to regions of the parameter space where the distance from the real data is lower. Figure A.13 shows the rapid convergence of the distance between the simulated and the real data.[86] We use the Manhattan norm distance, $\varrho(z(\mathbf{E},\omega)) = ||z(\mathbf{E},\omega)||_1 = \sum |z_i(\mathbf{E},\omega)|$. The results do not change significantly using different norms.

[86]Observe that the distance does not strictly decrease in the first 10,000 observations because the random component is generated at any iteration, thus the distance may increase if we keep the parameters constant.

A.4 Sensitivity Analysis with respect to network density and elasticity of network formation

We present here the sensitivity analysis with respect to the density of the network and the elasticity of network formation as measured by λ.[87] For both measures, we consider two different network topologies: the topology of the alumni connections, and of the Erdos-Renyi network. Table A.7 reports the $25th$, the median and the $75th$ percentile of the distribution of the bias, computed by subtracting the true value from our estimated posterior distribution for each parameter.

In the upper panel of the table, we explore the density of connections between nodes. For the network of alumni connections, we consider three cases: the high density network, which has the same density of the alumni network without any time restriction (about $d = 1.3$ percent); the medium density, which has the same density of the alumni network with 8 year restriction (about $d = 0.6$ percent); and the low density network, which has the same density of the alumni network with 4 year restriction (about $d = 0.3$ percent).[88] For the Erdos-Renyi network, we set $p = d$, keeping constant all of the other parameters. As before, for this exercise we also report the bias in the estimation of the parameters. These numbers show that network sparsity is not a necessary condition for the estimation of our model because the concentration of bias around zero does not appear to be related to network density.

In the lower panel, we study the performance of the model when the elasticity of network formation is changed in the alumni and Erdos-Renyi networks.[89] When $\lambda = 0$, the elasticity of network formation is zero and so model (12) is linear in $\theta_{i,j}$, as in standard spatial autoregressive models if $\theta_{i,j}$ is assumed to be the exogenous network. When $\lambda > 0$, and thus the elasticity of the network formation is positive, the model diverges from standard linear spatial autoregressive models because the social spillovers are nonlinear. We perform a simulation experiment to understand whether the performance of our estimation methodology varies when λ changes. We set $\lambda = 1, 2, 3$, and 4. The table presents the distribution of the estimation bias for the parameters for each of the respective values of λ. The results reveal no systematic pattern across values of λ, and that the distributions are mainly concentrated around zero for all values of λ, with similar dispersion. These results thus indicate that the performance of our methodology does not hinge on a particular value of λ.

[87] The density is measured as the ratio between the number of realized over the number of potential links, which is equal to $n(n-1)$ for a network with n nodes.

[88] The no year restriction means that two politicians are connected if they attended the same school, the 8 year restriction connects two politicians if they attended the same school within an interval of 8 years, and the 4 year restriction connects two politicians if they attended the same school within an interval of 4 years.

[89] The density is the lowest (about $d = 0.3$ percent) and all the other parameters are the same of the benchmark simulation used above and described in Section 8.5 in the Appendix.

A.5 Further Evidence on the Comparison between the Estimated and Observed Networks

To further analyze the differences between the estimated network and the actual ones, we report the densities of degree, closeness, clustering, and eigenvector centralities in Figure A.14-A.16 for each of the different networks. Interestingly, the density of the eigenvector centralities shows that our estimated network presents a marked bimodal distribution, which reveals the ability of our methodology to discriminate between more central and less central players. On the contrary, the seemingly normal distribution of centralities for the cosponsorship network seems compatible with a higher degree of randomness in the data generating process. The density of the closeness centrality of the estimated network is similar to the cosponsorship and committee networks, while it is concentrated on higher values than the one for the alumni network, reflecting the excessive sparseness of the connections in the alumni network. In terms of clustering and degree, the estimated network presents a smoother distribution than other networks, specifically with a higher number of nodes showing higher values of clustering and with more links than the alumni network.

Table A.8 more formally compares the estimated network with the cosponsorship, committee, and alumni networks. The table reports the mean across nodes for each network statistic, the T-statistics for equality of means, and its associated p-value. It also reports the Kolmogorov-Smirnov test statistic for the equality of the probability distributions. The results show that we can reject the hypothesis that the mean values of the centrality measures are the same in the estimated and actual networks in many cases, and that the Kolmogorov-Smirnov statistic always rejects the hypothesis that the empirical distribution of the centrality measure from our estimated network comes from the same distribution of any of the popular networks considered.

A.6 Data description - Further details

Volden and Wiseman [2014] identify nine factors that are important for legislative effectiveness. In our analysis, we include all of them as controls. In this section, we discuss each of them in turn.

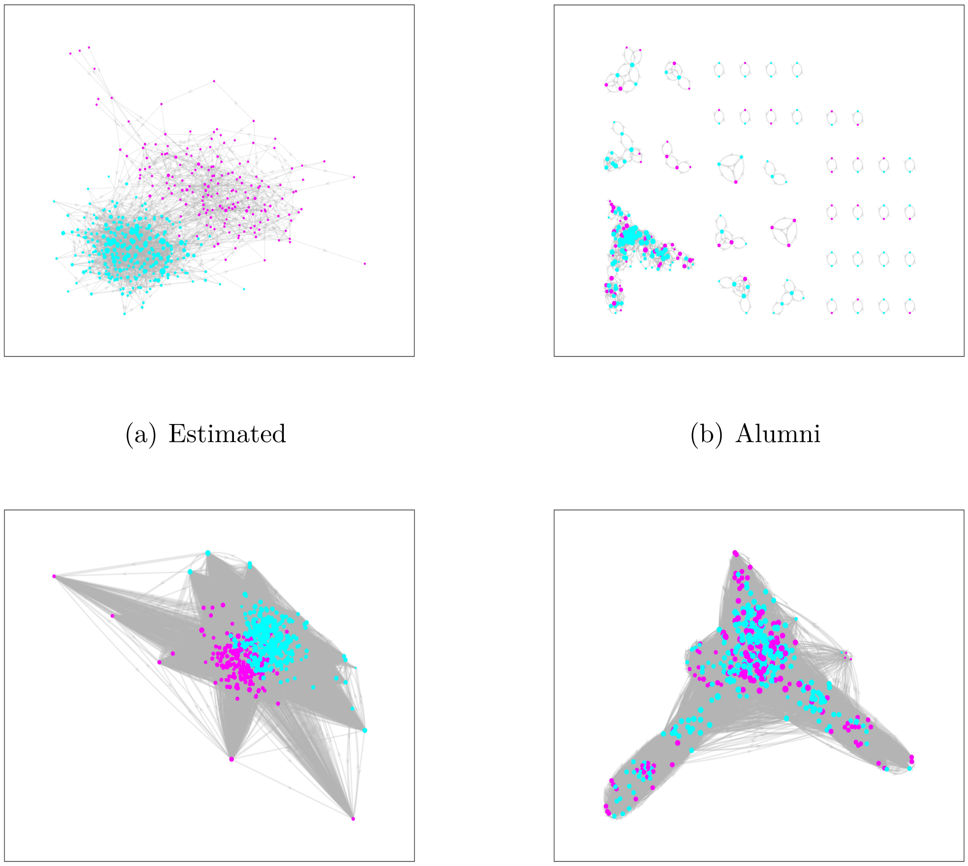
The first one is the number of years served as a member of Congress (*seniority*). As legislators spend more time in Congress, they are expected to become better and more effective at lawmaking. Consistent with the acquisition of skills over time, the second factor is previous *legislative experience*. Legislators who have previously served in state legislatures may be more effective than legislators without similar experience. Previous legislative experience is captured using a dummy taking a value of one if a legislator has previously served in a state legislature, and zero otherwise. It is then interacted with the state's level of professionalism, as measured by the index constructed by Squire [1992]. The next three factors (*party influence*, *committee influence*, and *legislative leadership*) capture the effect of institutional positions on the legislative process. Majority party members, committee chairs, members of the most powerful committees (Appropriation, Budget, Rules, and Ways and Means), and party leaders hold positions that may be associated with greater legislative effectiveness. The sixth factor captures *ideological considerations*. The Legislative Effectiveness Project data is merged with data from the Voteview Project.[90] Voteview provides data on legislators' ideological stance, as measured by the absolute value of the first dimension of the DW-nominate score created by McCarty et al. [1997]. A number of legislative politics studies suggest a negative correlation between this variable and legislative success, reflecting the idea that moderate policies obtain a larger consensus among the members of the House (see, e.g. Krehbiel [1992], Wiseman and Wright [2008]). The seventh factor includes the *demographic characteristics* of members of Congress. The experiences of women and legislators from other minority groups in terms of effective lawmaking are different from the average member of Congress, although the existing literature has not reached a consensus about the sign and the sources of these differences (Jeydel and Taylor [2003]; Volden and Wiseman [2014]; Volden et al. [2013]). The eighth factor captures natural coalition partners. Legislators from the same state may form a natural coalition, yielding greater legislative effectiveness. The *size of the congressional delegation*, which counts the number of districts in the state congressional delegation (and thus the number of Congress members in the House from the same state) may matter too. Legislators coming from larger congressional delegations may be more effective because they can find coalition partners among the members of their delegations. In contrast, the presence of more legislators interested in the same issues (the interests of the state) may result in a lower number of bills advanced in the legislative process for each legislator. The ninth factor is captured by the *degree of electoral competition*, as measured by the legislators' margin of victory (i.e. the percentage of total votes that separated the Congress member from the second-place finisher in the previous election). If voters value politicians' legislative effectiveness, then one would expect a positive

[90] See http://voteview.com.

relationship between legislators' levels of effectiveness and their margins of victory. The existence and sign of this relationship, however, is still a matter of debate. In fact, it is plausible to expect a negative correlation if electorally vulnerable legislators expend more energy to foster their agenda and increase support among voters. Alternatively, one may think that vulnerable legislators spend their energy on campaigning, while legislators in safe districts commit more time to the lawmaking process (see, e.g. Padro I Miquel and Snyder [2006], Volden and Wiseman [2014]).

A.7 Additional Figures

Figure A.1: ESTIMATED VS OBSERVED NETWORKS



(a) Estimated

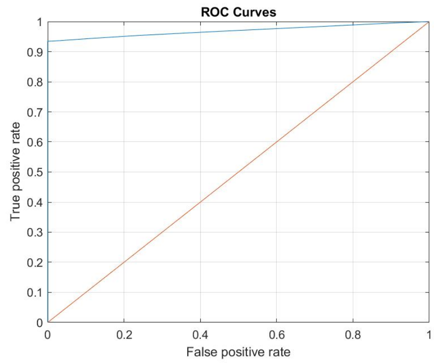
(b) Alumni

(c) Cosponsorship

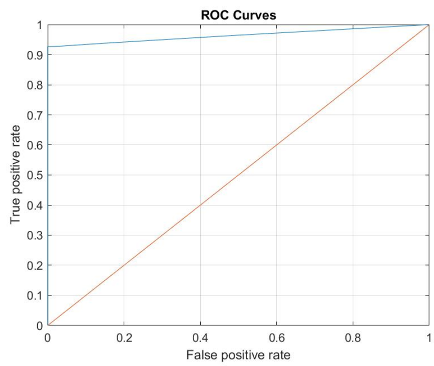
(d) Committee

NOTES. The estimated network is derived using the parameter estimates at the last iteration of the MCMCs for the 111th Congress. A dot is a politician. The color of the dot represents the party of the politician. Red nodes are Republicans. The networks are represented with force-directed layout with five iterations. It uses attractive forces between adjacent nodes and repulsive forces between distant nodes. For better visualization, the size of the nodes is equal to the (log) of their degree plus 2. The direct networks (cosponsorship and estimated) are transformed to indirect unweighted networks to have a clean comparison with the others. Given the direct network $D = \{d_{ij}\}$, its indirect unweighted counterpart is $U = \{u_{ij}\}$, where $u_{ij} = 1$ if d_{ij} or d_{ji} is different from zero, and zero otherwise. The alumni network is defined in Section 5.1. Cosponsorship activity is measured by directional links equal to one if j has cosponsored at least one bill proposed by i and zero otherwise. The ij_{th} element of the committee network is equal to the number of Congressional committees in which both i and j sit.

Figure A.2: NETWORK ESTIMATION
- GOODNESS OF FIT



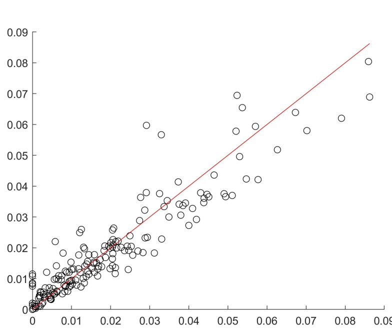
(a) Alumni network



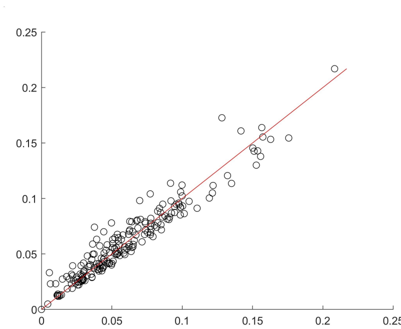
(b) Erdos-Renyi network

NOTES. Receiver Operating Characteristic (ROC) curve. The ROC curve is a plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied. For each threshold, the ROC curve reveals two ratios, the true positive rate TP/(TP + FN) and the false positive rate FP/(FP + TN), where TP is the number of true positives, FP is the number of false positives, TN is the the number of true negatives and FN is the number of false negatives. Y-axis: the true positive rate at various thresholds. X-axis: the false positive rate at various thresholds. The estimated network is derived using the parameter estimates at the last iteration of the MCMCs. The first of $\bar{r} = 5$ networks is represented.

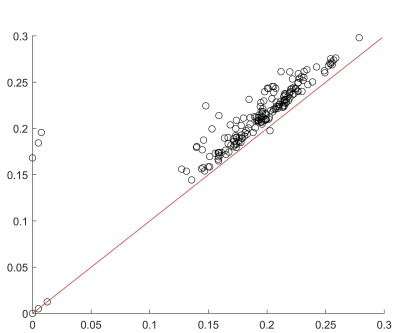
Figure A.3: NODE-LEVEL STATISTICS
- ESTIMATED VS TRUE NETWORK -
ERDOS-RENYI NETWORKS -



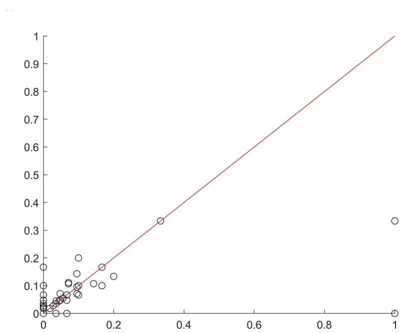
(a) Betweenness



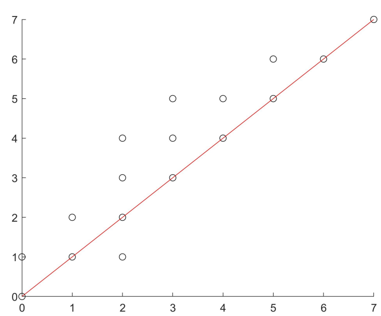
(b) Eigenvalue



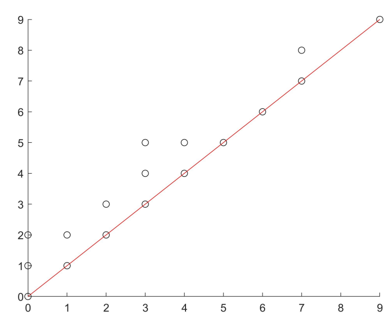
(c) Closeness



(d) Clustering



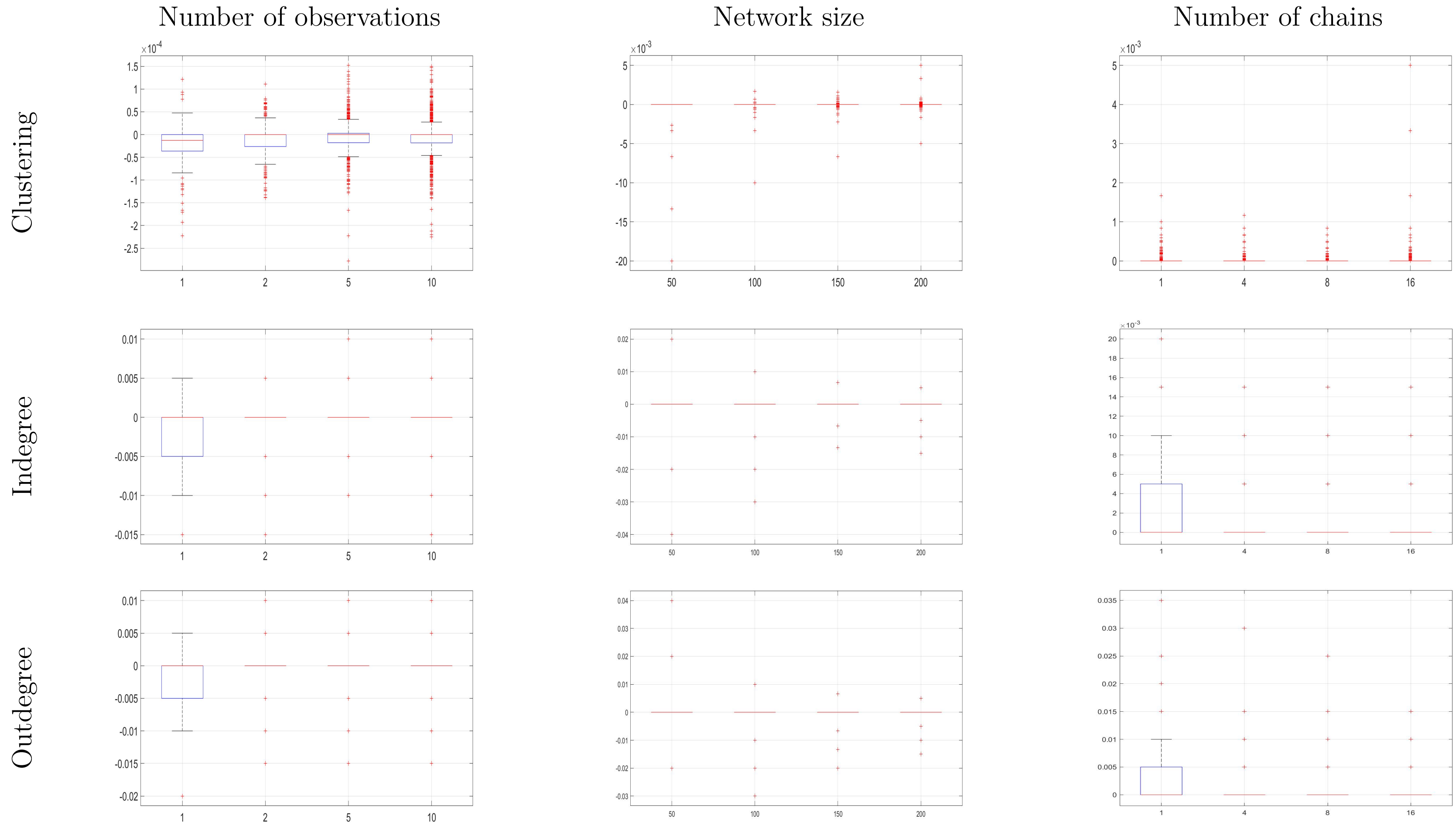
(e) Indegree



(f) Outdegree

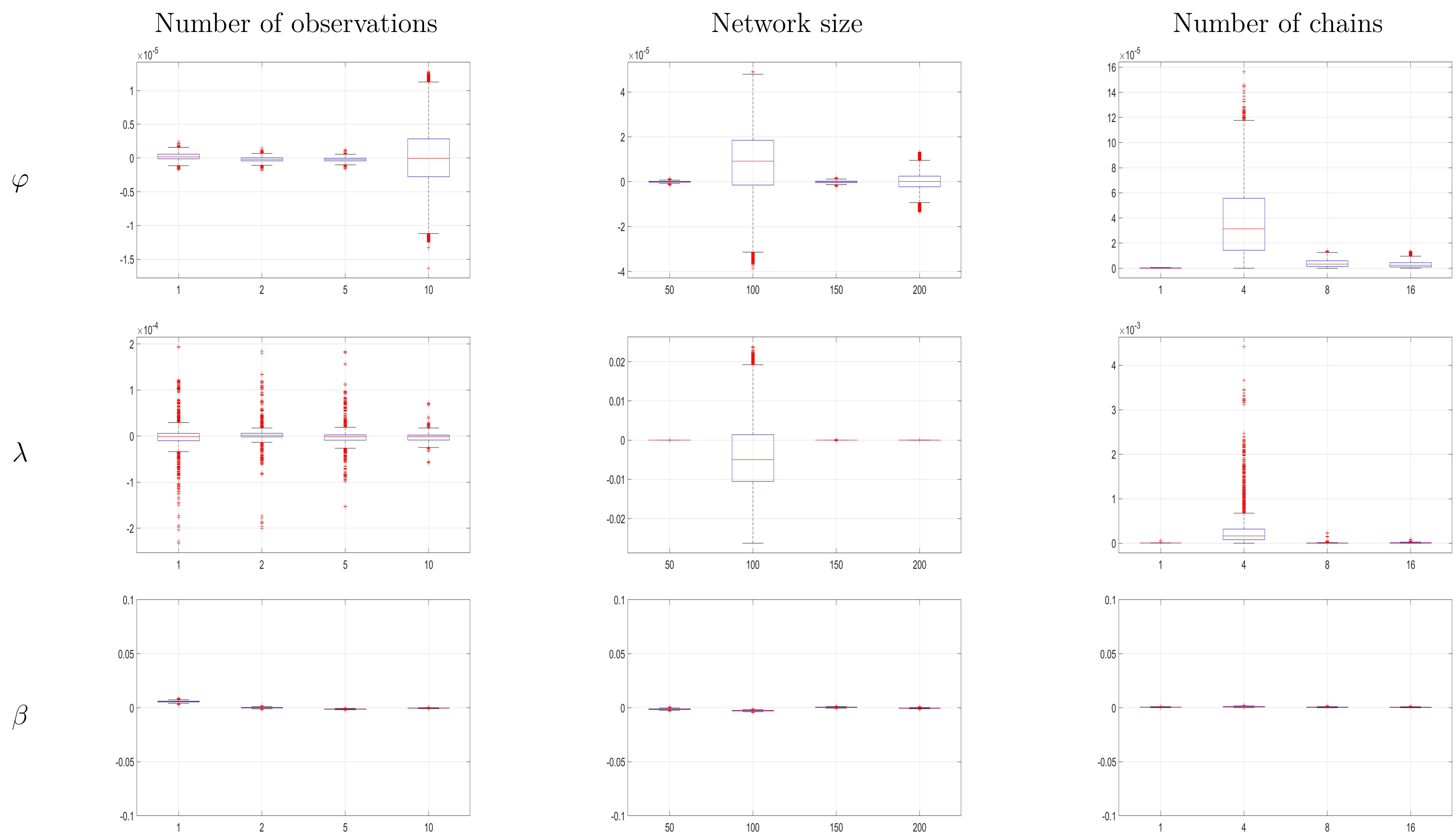
NOTES. X-axis: estimated value of node-level centralities as defined in Newman [2010]. Y-axis: true value of node-level statistic. The true values are the centralities of the true network in which the cost of forming a link depends on the Erdos-Renyi network. The estimated values are the centralities of the corresponding estimated network. See Section 8.5 for details on how the true network is constructed and estimated.

Figure A.4: ESTIMATION BIAS - NODE-LEVEL STATISTICS -



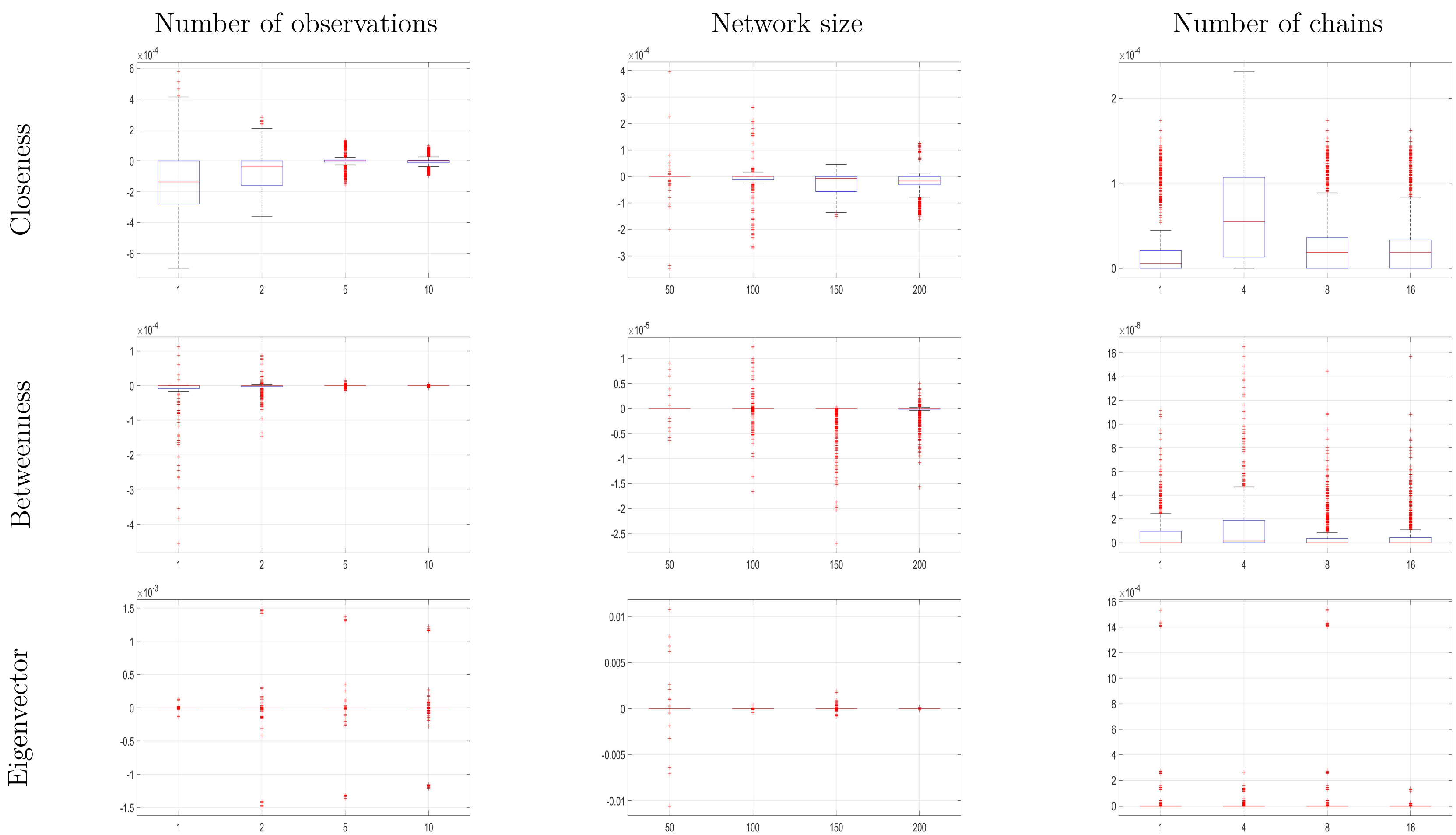
NOTES. X-axis: number of repeated observations of vector E (left panel); network size (center panel); number of MCMCs used for the estimation (right panel). Y-axis: distribution of the differences between estimated and true values of the centralities as defined in Newman [2010]. True and estimated values are constructed as described in Section 4.3.2. The DGP is described in detail in Section 8.5. The bottom and top edges of the boxes indicate the 25th and 75th percentiles of the distribution, respectively, and the central red mark indicates the median. The whiskers extend to the most extreme data points within 1.5 times the interquartile range. Values more than 1.5 times the interquartile range away from the top or bottom of the box (outliers) are plotted individually using the '+' symbol.

Figure A.5: ESTIMATION BIAS - PARAMETERS -
- ALUMNI NETWORK -



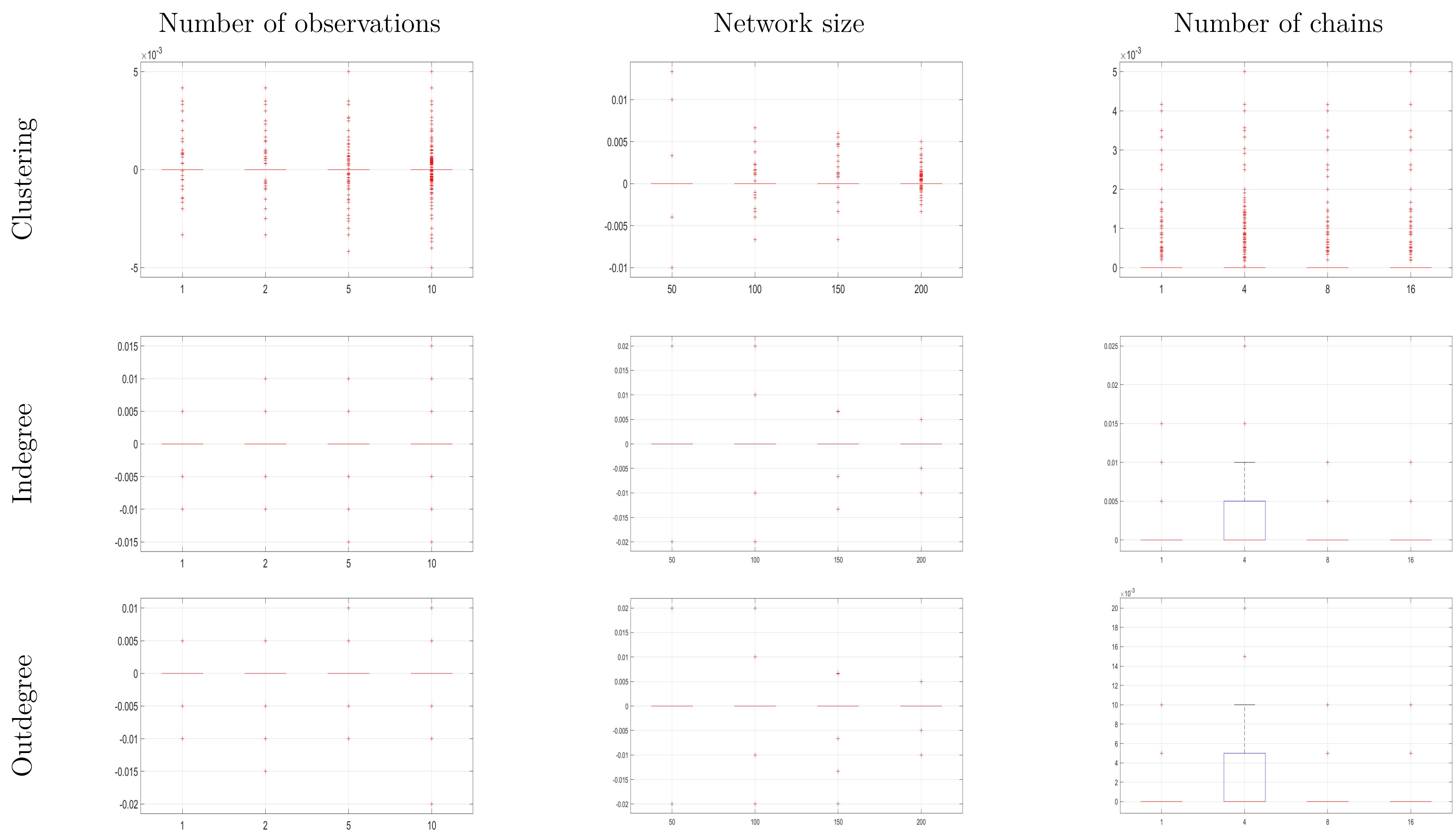
NOTES. X-axis: number of repeated observations of vector E (left panel); network size (center panel); number of MCMCs used for the estimation (right panel); Y-axis: distribution of the differences between the true values of the parameters and the estimated values. True and estimated values are constructed as described in Section 4.3.2. The parameters are defined as in equation (17). The DGP is described in detail in Section 8.5. The bottom and top edges of the boxes indicate the 25th and 75th percentiles of the distribution, respectively, and the central red mark indicates the median. The whiskers extend to the most extreme data points within 1.5 times the interquartile range. Values that are more than 1.5 times the interquartile range away from the top or bottom of the box (outliers) are plotted individually using the '+' symbol.

Figure A.6: ESTIMATION BIAS - NODE-LEVEL STATISTICS -
- ALUMNI NETWORK -



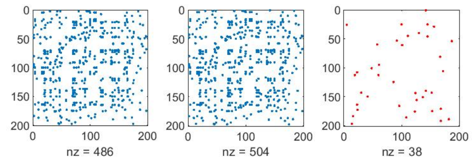
NOTES. X-axis: number of repeated observations of vector E (left panel); network size (center panel); number of MCMCs used for the estimation (right panel). Y-axis: distribution of the differences between estimated and true values of the centralities as defined in Newman [2010]. True and estimated values are constructed as described in Section 4.3.2. The DGP is described in detail in Section 8.5. The bottom and top edges of the boxes indicate the 25th and 75th percentiles of the distribution, respectively, and the central red mark indicates the median. The whiskers extend to the most extreme data points within 1.5 times the interquartile range. Values more than 1.5 times the interquartile range away from the top or bottom of the box (outliers) are plotted individually using the '+' symbol.

Figure A.7: ESTIMATION BIAS - NODE-LEVEL STATISTICS -
- ALUMNI NETWORK -

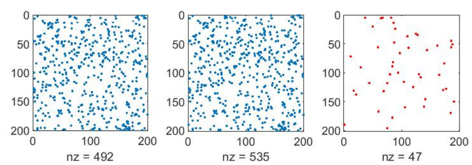


NOTES. X-axis: number of repeated observations of vector E (left panel); network size (center panel); number of MCMCs used for the estimation (right panel). Y-axis: distribution of the differences between estimated and true values of the centralities as defined in Newman [2010]. True and estimated values are constructed as described in Section 4.3.2. The DGP is described in detail in Section 8.5. The bottom and top edges of the boxes indicate the 25th and 75th percentiles of the distribution, respectively, and the central red mark indicates the median. The whiskers extend to the most extreme data points within 1.5 times the interquartile range. Values more than 1.5 times the interquartile range away from the top or bottom of the box (outliers) are plotted individually using the '+' symbol.

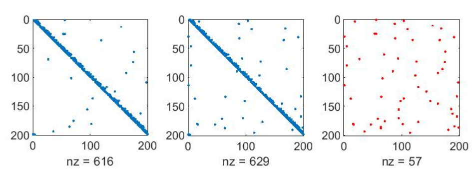
Figure A.8: ESTIMATED VS TRUE NETWORKS - DIFFERENT TOPOLOGIES -



(a) Alumni



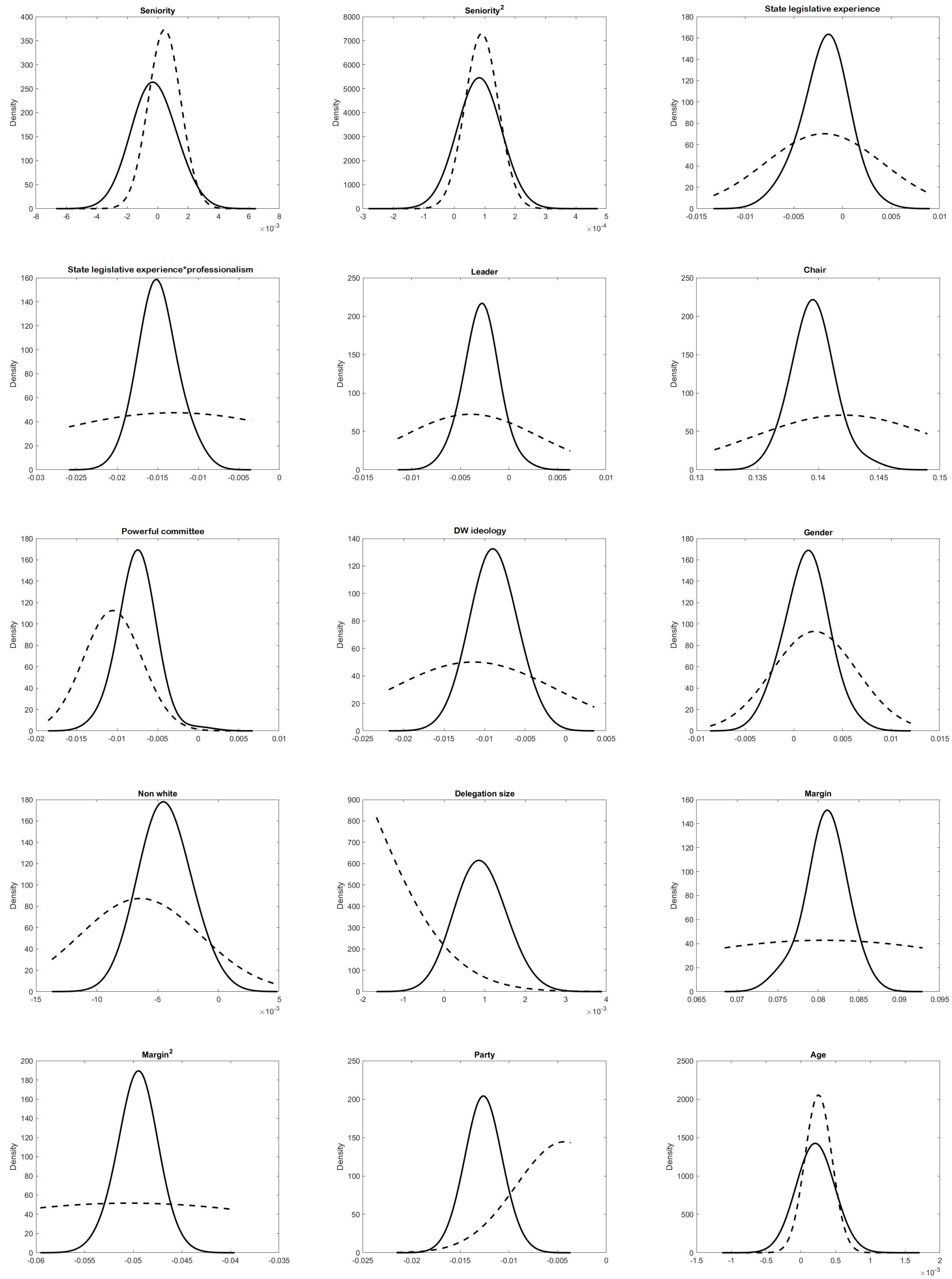
(b) Erdos-Renyi



(c) Circular

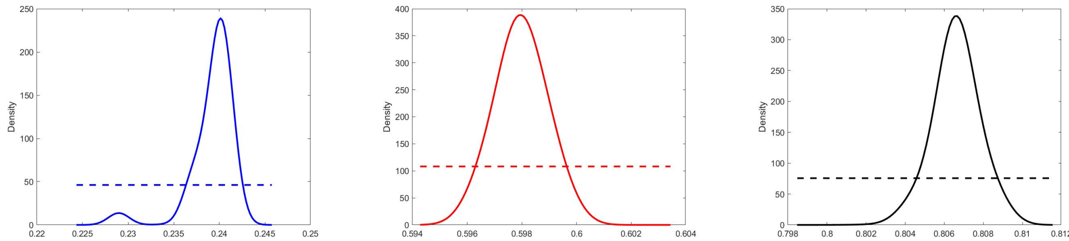
NOTES. Adjacency matrices of the true network, the estimated network (with blue dots) and their difference (with red dots) with n=200. The true and estimated networks are generated as described in Section 8.5. The DGP is described in detail in Section 8.5. The true networks in panels (a), (b) and (c) are generated, respectively, with alumni, Erdos-Renyi and Circular connections, as described in Section 8.5 and 4.3.2.

Figure A.9: ESTIMATED POSTERIOR DISTRIBUTIONS
- CONTROL VARIABLES -



NOTES. X-axis: parameter value, Y-axis: kernel density. The solid line represents the posterior distribution of the parameter estimated by the ABC algorithm, the dashed line depicts the prior distribution.

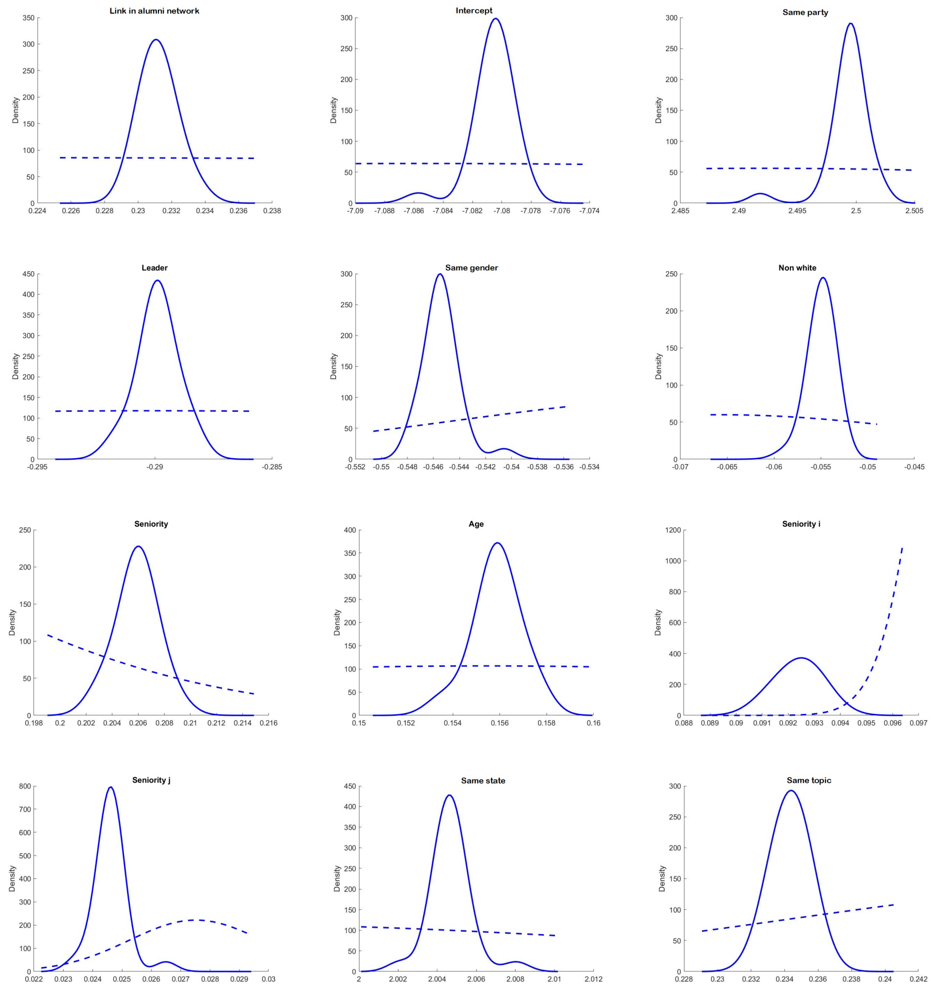
Figure A.10: ESTIMATED POSTERIOR DISTRIBUTIONS
- TARGET VARIABLES -



(a) ρ (b) λ (c) α

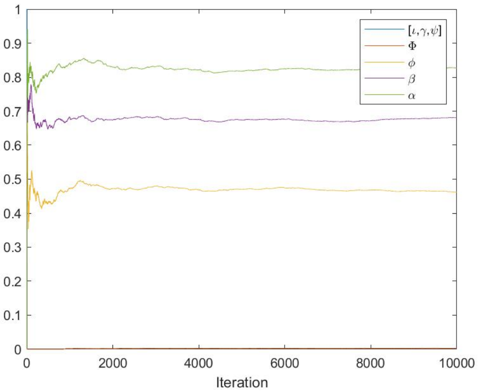
NOTES. X-axis: parameter value, Y-axis: kernel density. The solid line represents the posterior distribution of the parameter estimated by the ABC algorithm, the dashed line depicts the prior distribution.

Figure A.11: ESTIMATED POSTERIOR DISTRIBUTIONS
- LINK FORMATION -



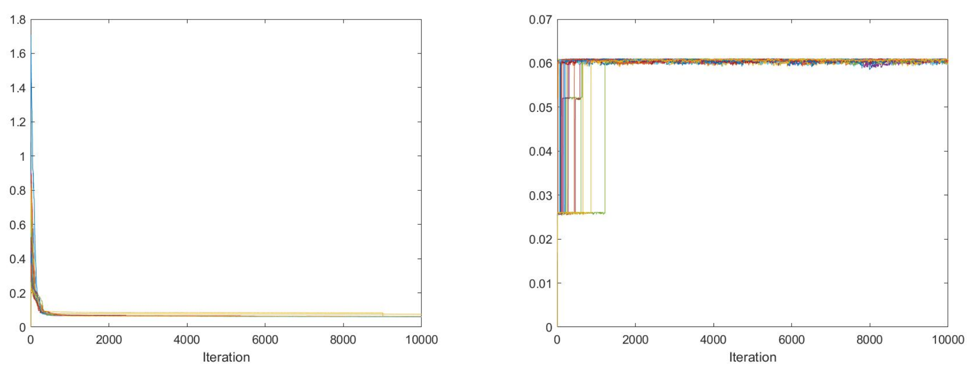
NOTES. X-axis: parameter value, Y-axis: kernel density. The solid line represents the posterior distribution of the parameter estimated by the ABC algorithm, the dashed line depicts the prior distribution.

Figure A.12: ACCEPTANCE RATE AT EACH ITERATION



NOTES. X-axis: MCMCs iteration in the second step of the ABC algorithm, Y-axis: acceptance rate. Acceptance rates of each parameter are averaged across the Markov chains.

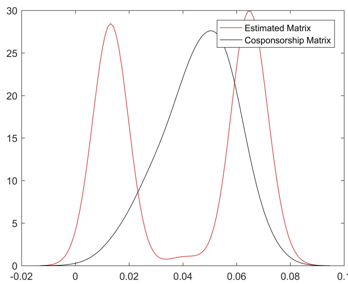
Figure A.13: DISTANCE BETWEEN SIMULATED AND REAL DATA AT EACH ITERATION



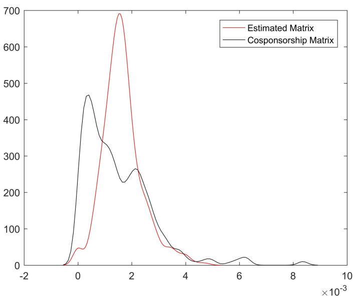
(a) First step (b) Second step

NOTES. X-axis: MCMCs iteration in the first step and second of the ABC algorithm, Y-axis: distance value at each iteration. Each line represents a Markov chain.

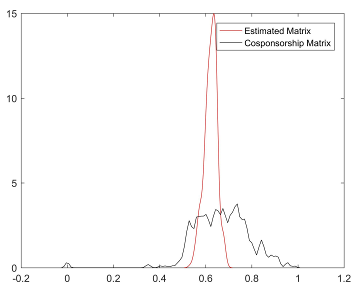
Figure A.14: ESTIMATED VS COSPONSORSHIP NETWORK
- DENSITIES OF NODE-LEVEL STATISTICS -



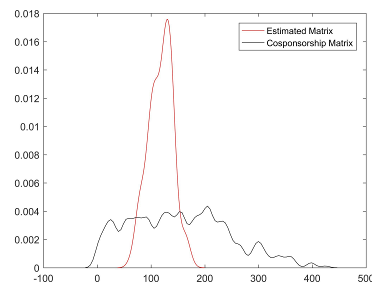
(a) Eigenvector



(b) Closeness

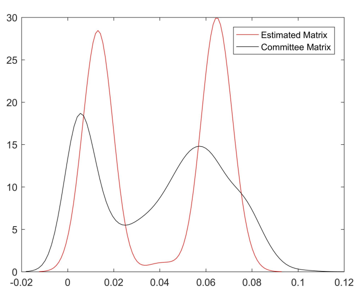


(c) Clustering

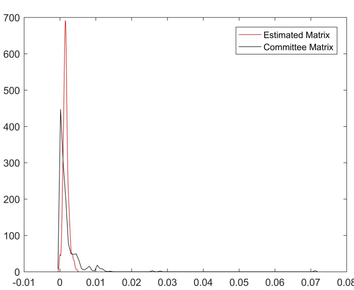


(d) Degree

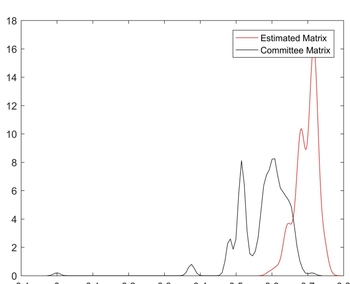
Figure A.15: ESTIMATED VS COMMITTEE NETWORKS
- DENSITIES OF NODE-LEVEL STATISTICS -



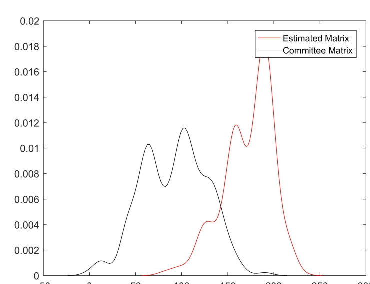
(a) Eigenvector



(b) Closeness

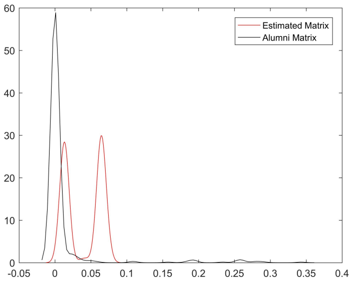


(c) Clustering

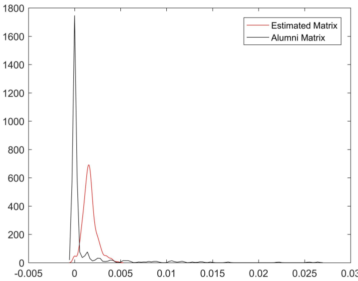


(d) Degree

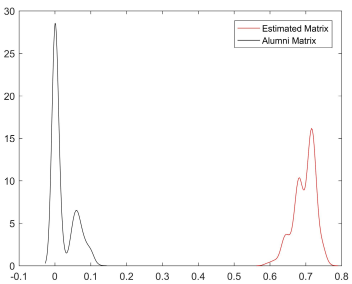
Figure A.16: ESTIMATED VS ALUMNI NETWORKS
- DENSITIES OF NODE-LEVEL STATISTICS -



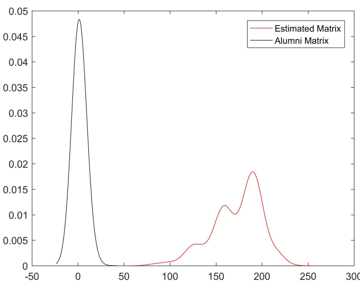
(a) Eigenvector



(b) Closeness



(c) Clustering



(d) Degree

NOTES. Kernel density estimate of node-level network measures. For each measure, the estimated network (in black) is compared with the observed network (in red). See Newman [2010] for the definition of network centrality measures. The estimated network is derived using the parameter estimates at the last iteration of the MCMCs for the 111th Congress. The alumni network is defined in Section 5.1. The ij_{th} element of the committee network is equal to the number of congressional committees in which both i and j sit. Cosponsorship activity is measured by directional links equal to one if j has cosponsored at least one bill proposed by i and zero otherwise. The direct networks (cosponsorship and estimated) are transformed to indirect unweighted networks to have a clean comparison with the others. Given the direct network $D = \{d_{ij}\}$, its indirect unweighted counterpart is $U = \{u_{ij}\}$, where $u_{ij} = 1$ if d_{ij} or d_{ji} is different from zero, and zero otherwise.

A.8 Additional Tables

Table A.1: NETWORK-LEVEL STATISTICS
- ESTIMATED VS TRUE NETWORKS -

	Estimated	True
Panel (a) - ALUMNI		
Density	0.0124	0.0139
Assortativity	7.0132	6.7424
Closeness	0.0595	0.0859
Betwenness	0.0488	0.0509
Degree	0.0738	0.0723
Clustering	0.7095	0.6482
Panel (b) - ERDOS-RENYI		
Density	0.0246	0.0266
Assortativity	0.4847	0.5152
Closeness	0.3125	0.3251
Betwenness	0.0457	0.0641
Degree	0.0360	0.0391
Clustering	0.0272	0.0287

NOTES. The true network is generated using equations (32)-(33) and an Erdos-Renyi network. The DGP is described in detail in Section 8.5. The estimated network is derived using the parameters' estimates at the last iteration of the MCMC. See Newman [2010] for the definition of network-level statistics.

Table A.2: SUMMARY STATISTICS

Variable name	Variable definition	Mean	Std
Party	Dummy variable taking value of one if the Congress member is a Democrat.	0.5060	0.5019
Gender	Dummy variable taking value of one if the Congress member is woman.	0.1723	0.3778
Non white	Dummy variable taking value of one if the member of Congress is African-American or Hispanic, and zero otherwise.	0.1388	0.3458
Seniority	Number of consecutive years in Congress.	5.7863	4.4388
Seniority^2	Number of consecutive years in Congress, squared.	53.1751	80.3864
DW ideology	Distance to the center in terms of ideology measured using the absolute value of the first dimension of the DW-nominate score created by McCarty et al. [1997].	0.5004	0.2236
Margin of victory	Margin of victory in the last election.	0.3526	0.2488
Margin of victory^2	Margin of victory in the last election, squared.	0.1862	0.2494
Committee chair	Dummy variable taking value of one if the Congress member is a chair of at least one committee.	0.0455	0.2084
Powerful committee	Dummy variable taking value of one if the Congress member is a member of a powerful committee (Appropriations, Budget, Rules, and Ways and Means).	0.2544	0.6355
Delegation size	Number of seats assigned to Congress member's state of election.	19.0988	15.4628
Leader	Dummy variable taking value of one if the member of Congress is a member of the party leadership, as reported by the Almanac of American Politics.	0.0496	0.2172
State legislative experience	Dummy variable taking value of one if the member of Congress served as a state legislator.	0.6260	0.6946
State legislative professionalism	State's level of professionalism [Squire, 1992].	0.1210	0.1779
Age	Age of the Congress member derived from the biographical files in `http://bioguide.congress.gov/`.	56.932	10.203
N. Obs.		2,176	

Source: Legislative Effectiveness Project (http://www.thelawmakers.org), Volden and Wiseman [2014] unless otherwise specified.

Table A.3: ESTIMATION RESULTS WITH UNOBSERVABLES

Dependent variable: Legislative Effectiveness Score (LES)		
	(1)	(2)
φ	0.0277 ***	0.0349 ***
	[1.0000]	[1.0000]
λ	0.5980 ***	0.0270 ***
	[1.0000]	[1.0000]
Party	-0.0124 ***	-0.0090 ***
	[0.0000]	[0.0000]
Gender	0.0012	0.0009
	[0.7295]	[0.7899]
Non white	-0.0042 ***	-0.0054 ***
	[0.0000]	[0.0000]
Seniority	-0.0001	-0.0006
	[0.4730]	[0.2085]
Seniority^2	0.0001 *	0.0001 **
	[0.9489]	[0.9555]
DW ideology	-0.0093 ***	-0.0126 ***
	[0.0000]	[0.0000]
Margin	0.0813 ***	0.0821 ***
	[1.0000]	[1.0000]
Margin^2	-0.0493 ***	-0.0514 ***
	[0.0000]	[0.0000]
Committee chair	0.1393 ***	0.1411 ***
	[1.0000]	[1.0000]
Powerful committee	-0.0083 **	-0.0101 ***
	[0.0417]	[0.0000]
Delegation size	0.0008 ***	0.0012 ***
	[0.9952]	[1.0000]
Leader	-0.0026 *	-0.0040 **
	[0.0820]	[0.0246]
State legislative experience	-0.0021	-0.0038 ***
	[0.1765]	[0.0000]
State legislative experience * State legislative professionalism	-0.0151 ***	-0.0136 ***
	[0.0000]	[0.0000]
Age	0.0002	0.0002
	[0.8861]	[0.8256]
$\sigma_{\epsilon,z}$		0.0002 ***
		[0.0000]
μ_1		0.0001 **
		[0.9525]
μ_2		0.0007
		[0.5768]
μ_3	-	-0.0006
		[0.3108]
μ_4	-	-0.0001
		[0.4833]
μ_5		0.0001
		[0.5691]
State fixed effects	Yes	Yes
Topic fixed effects	Yes	Yes
Congress fixed effects	Yes	Yes
N. Obs.	2,176	2,176

NOTES. Estimates of parameters in equation (17). The network formation model in column (1) is model (19). The network formation model in column (2) is model (28). In column (2), each parameter μ corresponds to the relative power of ϵ as η is generated with $\eta_{i,r} = \sum_{l=1}^{5} \mu_l \epsilon_{i,r}^{l}$. The median of the posterior distribution estimated with the ABC algorithm is reported for each coefficient. The empirical p-value of zero on the estimated posterior is reported in brackets. A precise definition of control variables can be found in Table A.2. *, **, and *** indicate statistical significance at the 10, 5 and 1 percent levels, based on empirical p-values.

Table A.4: COUNTERFACTUAL ANALYSIS -LEGISLATORS WITH EXTREME IDEOLOGIES-

		Female	Other	Seniority > sample mean	Other
Mean values of DW ideology					
Pre-treatment		0.4558	0.5096	0.4883	0.5124
Post-treatment	t				
	0.9	0.4498	0.4813	0.4663	0.4853
	0.8	0.4402	0.4547	0.4511	0.4533
	0.7	0.4147	0.4135	0.4203	0.4072
Share of connections to "Other"		94%	38%	71%	23%

NOTES. In the pre-treatment distribution, the averages are computed on observed data. In the post-treatment data distributions, the averages are computed on the transformed data, where the DW ideology of legislators above t are set equal to the mean below t. The connections to "Other" is the number of links that each category has with the relative "Other" category over the total number of links in the estimated network.

Table A.5: ESTIMATION RESULTS
- CONTEXTUAL EFFECTS -

	Endogenous network (1)	Endogenous network with exogenous contextuals (2)	
φ	0.0277 ***	0.0301 ***	
	[1.0000]	[1.0000]	
λ	0.5980 ***	0.5097 ***	
	[1.0000]	[1.0000]	
		direct (X)	contextual (HX)
Party	-0.0124 ***	-0.0049 ***	0.0004
	[0.0000]	[0.0000]	[0.6250]
Gender	0.0012	0.0019 ***	0.0004
	[0.7295]	[1.0000]	[0.6428]
Non white	-0.0042 ***	-0.0069 ***	-0.0002
	[0.0000]	[0.0000]	[0.4336]
Seniority	-0.0001	-0.0002	-0.0001
	[0.4730]	[0.4599]	[0.4914]
Seniority^2	0.0001 *	0.0001 *	0.0000
	[0.9489]	[0.9152]	[0.6606]
DW ideology	-0.0093 ***	-0.0112 ***	-0.0003
	[0.0000]	[0.0000]	[0.4030]
Margin	0.0813 ***	0.0801 ***	0.0001
	[1.0000]	[1.0000]	[0.5417]
Margin^2	-0.0493 ***	-0.0503 ***	-0.0004
	[0.0000]	[0.0000]	[0.4167]
Committee chair	0.1393 ***	0.1426 ***	0.0003
	[1.0000]	[1.0000]	[0.6487]
Powerful committee	-0.0083 **	-0.0105 ***	0.0001
	[0.0417]	[0.0000]	[0.5558]
Delegation size	0.0008 ***	0.0012 ***	0.0005 ***
	[0.9952]	[1.0000]	[1.0000]
Leader	-0.0026 *	-0.0037 ***	0.0001
	[0.0820]	[0.0000]	[0.5907]
State legislative experience	-0.0021	-0.0016	0.0000
	[0.1765]	[0.1038]	[0.5101]
State legislative experience *	-0.0151 ***	-0.0128 ***	-0.0001
State legislative professionalism	[0.0000]	[0.0000]	[0.4763]
Age	0.0002	-0.0008 ***	-0.0002 ***
	[0.8861]	[0.0000]	[0.0000]
State fixed effects	Yes	Yes	
Topic fixed effects	Yes	Yes	
Congress fixed effects	Yes	Yes	
State and topic contextual effects	No	Yes	
N.Obs.	2,176	2,176	

NOTES. Estimates of parameters in equation (17). In column (1) the model is estimated without contextual effects. In column (2) the model is estimated with contextual effects as described in Section 6.4. The median of the posterior distribution estimated with the ABC algorithm is reported for each coefficient. The empirical p-value of zero on the estimated posterior is reported in brackets. A precise definition of control variables can be found in Table A.2. *, **, and *** indicate statistical significance at the 10, 5 and 1 percent levels, based on empirical p-values.

Table A.6: LINK FORMATION
- CONTEXTUAL EFFECTS -

Dependent variable: probability of forming a link

	(1)	(2)
Link in alumni network	0.2310 ***	0.1797 ***
	(1.0000)	(1.0000)
Seniority [1 = same quartile]	0.2060 ***	0.1548 ***
	(1.0000)	(1.0000)
Seniority i	0.0924 ***	0.0959 ***
	(1.0000)	(1.0000)
Seniority j	0.0246 ***	0.0276 ***
	(1.0000)	(1.0000)
Same state [1 = yes]	2.0048 ***	1.9900 ***
	(1.0000)	(1.0000)
Same topic [1 = yes]	0.2344 ***	0.2165 ***
	(1.0000)	(1.0000)
Leader [1 = both leaders]	-0.2899 ***	-0.2672 ***
	(0.0000)	(0.0000)
Same gender [1 = yes]	-0.5456 ***	-0.5329 ***
	(0.0000)	(0.0000)
Same race [1 = both white or both non white]	-0.0547 ***	-0.0825 ***
	(0.0000)	(0.0000)
Same party [1 = yes]	2.4994 ***	2.5234 ***
	(1.0000)	(1.0000)
Age [1 = same quartile]	0.1559 ***	0.1532 ***
	(1.0000)	(1.0000)
Intercept	-7.0805 ***	-7.1600 ***
	(0.0000)	(0.0000)
N.Obs.	2,176	2,176

NOTES. Estimates of parameters in equation (19) are reported. In column (1) the endogenous network model is estimated without exogenous contextual effects. In column (2) the endogenous network model is estimated with exogenous contextual effects as detailed in Section 6.4. The median of the posterior distribution estimated with the ABC algorithm is reported for each coefficient. The empirical p-value of zero on the estimated posterior is reported in brackets. Seniority i and Seniority j denote the seniority of legislator i and j, respectively. The rest of the independent variables are dummies capturing differences in characteristics between i and j. A precise definition of the variables at the individual level can be found in Table A.2. The threshold for unobservables is equal to one standard deviation above the mean of their distribution. *, **, and *** indicate statistical significance at the 10, 5 and 1 percent levels, based on empirical p-values.

Table A.7: ESTIMATION BIAS
- NETWORK DENSITY AND ELASTICITY OF NETWORK FORMATION -

Parameter	$\hat{\varphi}$			$\hat{\lambda}$			$\hat{\beta}$		
Percentiles	*25*	*50*	*75*	*25*	*50*	*75*	*25*	*50*	*75*
Network density (d)									
Alumni									
Low	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0010	0.0012	0.0013
Medium	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0016	-0.0014	-0.0012
High	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0005	-0.0004	-0.0002
Erdos-Renyi									
Low	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0015	-0.0014	-0.0011
Medium	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0012	0.0013	0.0015
High	0.0013	0.0016	0.0020	0.0013	0.0016	0.0020	0.0005	0.0007	0.0009
λ									
Alumni									
4	-0.3606	-0.2208	0.1872	-0.3606	-0.2208	0.1872	-0.0005	-0.0003	-0.0002
3	-0.5777	-0.4916	-0.2020	-0.5777	-0.4916	-0.2020	-0.0020	-0.0016	-0.0012
2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0005	-0.0004	-0.0002
1	-0.0035	-0.0033	-0.0031	-0.0035	-0.0033	-0.0031	0.0021	0.0022	0.0024
Erdos-Renyi									
4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0009	0.0010	0.0012
3	-0.0032	-0.0028	-0.0022	-0.0032	-0.0028	-0.0022	-0.0011	-0.0009	-0.0007
2	0.0013	0.0016	0.0020	0.0013	0.0016	0.0020	0.0005	0.0007	0.0009
1	-0.0004	-0.0001	0.0003	-0.0004	-0.0001	0.0003	0.0029	0.0031	0.0033

NOTES. The DGP is described in detail in Section 8.5. The true values of the parameters are fixed and generated using equations (32)-(33). The connections are generated from an alumni network and an Erdos-Renyi network, as defined in Section 4.3.1. The estimated values are taken from the posterior distribution of 16 MCMCs in the ABC algorithm. The 25^{th}, 50^{th}, and 75^{th} percentiles from the distribution of the differences between estimated and true values in the MCMCs after a burning period of 10,000 iterations are reported. The high density network has the density of the alumni network without restrictions, $d = 1.3\%$; the medium density network has the density of the alumni with an 8 year restriction, $d = 0.6\%$; the low density network has the density if the alumni with a 4 year restriction, $d = 0.3\%$.

Table A.8: NETWORK DIFFERENCES - STATISTICAL TESTS

	Estimated Mean	Mean	T-stat	p-value	Kolmogorov Smirnov test	p-value
Cosponsorship						
Indegree	4.1319	154.4508	-1.6008	0.0548	0.9472	0.0000
Clustering	0.0446	0.6825	-7.0556	0.0000	0.9982	0.0000
Between	0.0003	0.0001	0.7726	0.7801	0.5593	0.0000
Closeness	0.0448	0.1346	-3.4155	0.0003	0.9936	0.0000
Eigenvector	0.0068	0.0094	-0.0850	0.4661	0.2256	0.0000
Committee						
Indegree	8.1250	87.2840	-2.3085	0.0105	0.9793	0.0000
Clustering	0.0446	0.7071	-3.0494	0.0012	0.9908	0.0000
Between	0.0003	0.0001	0.6108	0.7293	0.4770	0.0000
Closeness	0.0660	0.1142	-2.7269	0.0032	0.9835	0.0000
Eigenvector	0.0068	0.0084	-0.0518	0.4794	0.3203	0.0000
Alumni						
Indegree	8.1250	1.3805	1.3106	0.9049	0.7753	0.0000
Clustering	0.0446	0.1684	-0.3673	0.3567	0.3157	0.0000
Between	0.0003	0.0000	0.8327	0.7974	0.7904	0.0000
Closeness	0.0660	0.0027	7.0696	1.0000	0.9972	0.0000
Eigenvector	0.0068	0.0021	0.1763	0.5700	0.5386	0.0000

NOTES. Node-level statistics are considered. See Newman [2010] for the definition of network-level statistics. The first four columns test differences in means, the last two columns test the difference between the two distributions. The alumni network is defined in Section 5.1. Cosponsorship activity is measured by directional links equal to one if j has cosponsored at least one bill proposed by i, and zero otherwise. The ij_{th} element of the Committee network is equal to the number of Congressional committees in which both i and j sit. The direct networks (cosponsorship and estimated) are transformed to indirect unweighted networks to have a clean comparison with the others. Given the direct network $D = \{d_{ij}\}$, its indirect unweighted counterpart is $U = \{u_{ij}\}$, where $u_{ij} = 1$ if d_{ij} or d_{ji} is different from zero, and zero otherwise.