Endogenous social interactions with unobserved networks*

Marco Battaglini[†], Eleonora Patacchini[‡], Edoardo Rainone[§]

March 3, 2021

Abstract

We present a model of endogenous network formation to recover unobserved social networks using only observable outcomes. We propose a novel equilibrium concept that allows for a sharp characterization of equilibrium behavior and that yields a unique prediction under testable conditions. While the equilibrium is characterized by a large number of nonlinear equations, we show that it can be efficiently employed to recover the networks by an appropriately designed approximate Bayesian computation method. We apply the model to recover the network of social links between lawmakers in the U.S. Congress using data from the 109th to 113th legislatures. We show that social connections are important for legislators' productivities and we identify some of the key determinants of network centralities in the U.S. Congress. JEL: C31, D85, D72

1 Introduction

Social networks are widely assumed to play a key role in many economically relevant environments, but they are rarely directly observed. Social networks, for example, are considered important in the determination of adolescents' risky behavior and educational achievements, in the determination of lawmakers' careers and effectiveness in legislatures, and in the diffusion of investment choices and business practices among CEOs and directors of corporate boards.[1] Such personal relationships, however, are rarely formally documented and indeed often intentionally kept private. The network

*For useful comments and discussions we thank Coralio Ballester, Vincent Boucher, Antonio Cabrales, Chih-Sheng Hsieh, Valerio Leone Sciabolazza Angelo Mele, Pedro CL Souza, Mathieu Taschereau-Domouchel, seminar participants at the Kellogg School of Business, the Einaudi Institute for Economics and Finance, the plenary lecture of the 2018 Public Economic Theory conference (PET 2018), and the 2018 Barcelona GSE Summer Forum. We also thank Neelanjan Datta and Julien Neves for excellent research assistance. The views expressed here do not necessarily reflect those of the Bank of Italy.

†Cornell University and EIEF E-Mail: battaglini@cornell.edu

‡Cornell University E-Mail: patacchini@cornell.edu

§Bank of Italy E-Mail: edoardo.rainone@bancaditalia.it

[1] We discuss these applications in Section 2.2.

of interbank credit exposures is an important determinant of the fragility of a financial system. Yet again these links are not directly observable even by professional investors, who need to reconstruct them from aggregate accounting data or partial information on off-balance sheet positions. In all these scenarios, the traditional approach in the literature has been to assume that latent social networks can be approximated by some observable proxies, such as having attended the same school or some measure of physical proximity, and then to assess the impact of these proxies on observable outcomes.

In this paper, we pose the reverse question: can we use observable outcomes to recover social networks and to assess their economic and social relevance? The challenge in this exercise is that a network of social connections between n agents is on the order of n^2, while a vector of observations is of order n. Even when repeated observations are available, it is rare to observe the outcome associated with a social network for more than a few periods: repeated observations cannot fully compensate for the difference in dimensionality. We propose a new theory of network formation in which both the choice of forming social links among the players and their activity in the network are endogenous. The theory is applicable to complex networks and, under testable conditions, it gives us a unique prediction for the network and the observable outcomes, thus allowing us to reduce the dimension of the problem to the parameters of the model. We use our theory to structurally estimate social links in the 109th to 113th U.S. Congresses using data on lawmakers' characteristics and on their legislative effectiveness.

Our theory has two stages. In the first stage, the players target their efforts to form social links with specific other players. The effort required to form social links is costly, and bilateral links between two players may depend on both of the players' efforts, according to a given production function. In the second stage, the players' effectiveness depends on their efforts and the effectiveness of the players with whom links have been established in the previous stage. The preferences of the players in the linking process may not only depend on their observable characteristics, but also on unobservable factors that may be correlated with variables that directly affect the players' effectiveness.

We are able to overcome the complications in solving and estimating the model described above because of two methodological contributions. First, we introduce a new equilibrium concept that we call *Network Competitive Equilibrium* (hereafter, NCE). The difficulty in solving the game described above is that, by establishing a link, a legislator generates direct effects on his/her effectiveness, but also a cascade of indirect spillover effects that complicates the analysis: legislator i's change in effectiveness after linking to j affects the effectiveness of all the others directly or indirectly linked to i, including j. These complications are similar to the problems that arise when studying a general equilibrium in an exchange economy. In such an environment, a change in an agent's demand has a direct obvious effect on an agent's utility and a cascade of indirect effects on equilibrium prices, the budget of other agents, and their actions. In competitive analysis, this problem is solved by assuming that agents are "price takers:" agents solve their optimization program taking prices as given. Prices, however, must clear the market in equilibrium. Such analysis is motivated by the observation that, in many exchange economies, each agent only has a marginal impact on equilibrium prices, thus allowing them to ignore the indirect effects. Similarly, in our approach, players choose their socialization efforts taking the other players' equilibrium effectivenesses as given; these effectiveness levels, however, need to be consistent with individual choices in equilibrium. We show that a NCE can be characterized by a system of nonlinear equations.

Our second methodological contribution is how we use the analytical characterization of the equilibrium conditions to estimate the model by Bayesian methods. Because our characterization makes it impossible to state an analytic likelihood function, we estimate the model by an *approximate Bayesian computation* method (hereafter, ABC), a computational approach that has proven useful in population genetics and other applications that require large scale models.[2]

Using Monte Carlo simulations, we systematically explore the performance of our estimation technique by studying environments in which we change the number of nodes, the number of periods in which behavior is observed, and other characteristics of the network topology. We show that our approach performs well for networks with at least 150 members, even if the networks are dense and we only observe outcomes in a few periods.[3]

The assumption behind the NCE that players are "price takers" makes our approach appropriate for applications with large networks in which the status of any single agent depends on the linking decision of many other agents, though it may be a limitation in other environments. Our approach, however, can be extended to environments in which there are both small "non-atomistic" players (i.e., the followers) and large "atomistic" players who may strategically affect the others (i.e., the leaders). We show that in these cases, the NCE can be used to solve for the connections of the followers, thus allowing us to directly estimate the connections of the larger players.

We apply our methodology to estimate the social network in the U.S. Congress using data from the 109th to 113th Congresses. We find evidence that social connections affect legislative effectiveness. We estimate that a one percent increase in the social connectedness of a legislator i (a measure of the effectiveness of the other legislators connected to i that we will define precisely) induces a 0.80 percent increase in individual effectiveness. Consistent with the endogeneity of the network, we also find that the elasticity of link formation with respect to other legislators' effectiveness levels is significantly positive.

Perhaps more importantly, the estimation of the social network gives us insight into the determinants of social connections. Consistent with the existing literature on congressional politics, party affiliation is the most significant factor determining social connections, with Republicans more linked to Republicans and Democrats more linked to Democrats. More surprisingly, we find

[2]See Rubin [1984] and Marjoram et al. [2003] for a discussion of ABC methods and their applications.

[3]Indeed, one period is sufficient, though the estimates improve with three to five periods.

that interpersonal relations formed long before being elected to Congress are another a key factor, which is about one-tenth as large as party affiliation. This is in line with evidence that alumni connections predict cosponsorships and voting behavior in Congress (Cohen and Malloy [2014] and Battaglini and Patacchini [2018]). We also find that weak links across political lines are important, an observation long suggested in the sociological and political science literature (see Granovetter [1973] and, more recently, Kirkland [2011]). Finally, the model allows us to estimate social network effects without having to rely on a specific, observable adjacency matrix as an approximation of social connections. For example, by using exclusively alumni networks as a proxy (as done by Cohen and Malloy [2014] or Battaglini and Patacchini [2018]), we ignore the importance of party affiliation; and by using only cosponsorships (as done by Fowler [2006], for example), we ignore the importance of alumni connections.

We show that both party affiliation and alumni connections are important ingredients to determine the social connections that matter for productivity in the U.S. Congress.

We use our structural model in two counterfactual exercises. In the first, we consider the effects of a policy that mitigates the ideology of the most extreme lawmakers, and, in the second, we consider a policy that mitigates the legacy of "old boy" networks as perpetrated by the alumni connections. As we discuss in more detail below, these mitigation policies are likely to be the result of campaign finance or electoral reforms that broaden the pool of candidates that can afford to run for Congress. The analysis provides a number of new insights on the relationship between social ties, power, and effectiveness in the U.S. Congress. When we moderate the ideology of the most extreme lawmakers, the model shows an overall improvement in productivity, but there are also *redistributive effects*: women, in particular, tend to become more effective. Another interesting redistributive effect is observed when we remove the influence of the alumni connections: women and lawmakers from minority groups see their centrality increase in the equilibrium social network. When we eliminate the influence of the alumni connections, we also predict that more senior legislators, chairs of committees, and Democrats gain centrality. This could be explained by the fact that links based on alumni ties partly crowd out links with legislators with higher experience and/or influence: after all, legislator have a given budget of time to socialize, so favoring one type of link may come at the cost of other potentially useful links.

The remainder of this paper is organized as follows. Section 2 presents our model of legislative behavior and the formation of social connections. Section 3 defines and characterizes our equilibrium concept, the NCE. Section 4 defines the econometric specification of the model, discusses the estimation method, and presents a set of simulations to explore the performance of our approach in finite samples and as we change the estimation environment. Section 5 estimates the model using data from the 109th to 113th U.S Congresses and presents a set of counterfactual exercises. Section 6 presents a few extensions to the model and discusses the robustness of the results. Section 7 concludes. In the reminder of this section, we discuss the related literature.

Related literature There is an important economic literature on the estimation of social networks. The traditional approach is to assume that the relevant social network is *observable*, and that it can be interpreted as a realization from unknown, latent data generating processes. This allows researchers to exploit features of the architecture of the observed network (e.g., the frequency of transitive triplets, the diameter, average path length, and the average clustering coefficient) to recover statistical determinants of the links.[4]

The literature has only recently started to address the issue of estimating an *unobservable* network using other observable economic outcomes. This approach involves taking the network as exogenous and stable over time so that repeated observations from the same network can be used in the estimation. The complication here is that the number of network connections is much higher than the number of economic outcomes that can be typically used for the estimation.[5] De Paula et al. [2018], Rose [2018] and Battaglini et al. [2020a] use high-dimensional estimation techniques to estimate social networks, which can bypass the dimensionality problem when the networks are sufficiently sparse (relative to the number of observations).[6] While this and other related approaches are versatile and can be applied to many environments with minor changes, they have two main limitations.[7] First, they typically require assuming an exogenous linear model of behavior and thus do not allow for the endogeneity of the network. Second, and more importantly, they require assuming that a social network is sparse and fixed over the repeated observations and/or the vector of outcomes used for the estimation can be observed for many sessions. These assumptions are problematic for environments such as ours: as we will argue below, legislative networks are generally dense and cannot be observed repeatedly because elections are held every few years.[8] Instead, our approach relies on the structure generated by our endogenous model of link formation to reduce the dimensionality of the parameters to be estimated. We then use the observable outcome from just one or a few Congresses to structurally estimate the parameters of the model.

To our knowledge, we are the first to model the network formation as a "competitive equilibrium" as described above. Two very different approaches have been adopted to model complex network formation processes for empirical analysis. The first approach is based on stochastic best response dynamics, in which myopic agents form random links sequentially, thus forming a Markov

[4]Recent contributions following this empirical approach are, among others, Christakis et al. [2010], Mele [2017], and Badev [2017], Liu et al. [2012], Konig [2016], and Boucher [2020].

[5]With n legislators observed for T sessions, the first is on the order of $n(n-1)$ ignoring connections to selves, the second just of order $n \cdot T$ with $T << n$. In the case of legislative networks $n \approx 400$, while T, depending on the frequency of observation, ranges from 2 if the outcome is measured per session, to 24 if measured montly.

[6]A spatial econometrics model with an unobserved and stochastic network is also proposed by Souza [2014]. Breza et al. [2017] propose a method to estimate social links using aggregated relational data.

[7]For a related approach, see Manresa [2016] who, however, only considers exogenous peer effects.

[8]Elections for the U.S. Congress are held every two years. While reelection rates are very high (about 90-95 percent of members are reelected), the average length of service is about five terms in the U.S. House of Representatives as of January 2013; and, on average, over 10 percent of first-term members do not seek reelection. It is therefore problematic to assume social networks in Congress persist for more than a few Congresses.

chain of networks. In this approach, an observed network is interpreted as a realization from the stationary distribution of the Markov process, which is obtained by simulation.[9] The second approach is to study networks that satisfy pairwise stability, a stability concept introduced by Jackson and Wolinsky [1996]. This approach typically generates a large set of equilibrium predictions and thus leads only to partial identification under strong sparsity conditions on the network.[10] In Section 3.1, we discuss the relative advantages and limitations of these other approaches in comparison to our approach, after we have described our model and equilibrium solution in greater detail.

Closer to our approach in modeling network formation are the works by Battaglini et al. [2020b] and Acemoglu and Azar [2018]. The first paper also looks at the effectiveness of legislators: as in our model, legislators choose effort and take the centrality of the other players as given in this model. The legislators' network, however, is taken as exogenous and not modeled. The second paper studies the endogenous choices of firm input combinations in a competitive environment. In this model, marginal productivities of inputs depend on the entire input/output network, but are taken as given by firms when choosing suppliers. Contrary to us, however, the authors assume that the production network is observable and therefore do not attempt to estimate it: their focus is on studying its impact on aggregate productivity in a dynamic model of growth.

In terms of the main application of our methodology, our work relates to a recent literature studying the impact of social networks on legislative behavior in Congress.[11] While scholars in political science have recognized the relevance of social connections between lawmakers for quite some time, only recently have data availability and advances in network analysis allowed us to move beyond descriptive analyses. Previous to our work, the role of social connections on legislative effectiveness has been studied by Fowler [2006], Kirkland [2011], and Battaglini et al. [2020b], among others. Fowler [2006] was among the first to document a relationship between effectiveness and measures of centrality in the cosponsorship network of the U.S. Congress.[12] Kirkland [2011], using data on the cosponsorship networks at the level of U.S. state legislatures, confirmed Fowler's results, emphasizing the importance of distinguishing between weak ties among legislators. These papers, however, have not addressed the issue that both the cosponsorship network and the legislators' levels of effectiveness are endogenous; and that the cosponsorship network is only a proxy for the true unobserved social connections among lawmakers. Battaglini et al. [2020b] presents a first attempt to control for endogeneity of the cosponsorship network by

[9]Using this approach, Christakis et al. [2010], Mele [2017], and Badev [2017] provide microfoundations for the exponential random graph approach. Jackson and Rogers [2007], Liu et al. [2012], Konig [2016], and Boucher [2020] have characterized the distribution of networks emerging from alternative sequential models.

[10]Miyauchi [2016], De Paula et al. [2018b], and Sheng [2018] each provide a partial identification analysis of network formation models based on the cooperative solution concept of pairwise stability under restrictions on the complexity of network connections.

[11]See Victor et al. [2016] and Battaglini and Patacchini [2019] for recent surveys.

[12]The cosponsorship network is the network in which a link from i to j is established if j has cosponsored bills by i.

using the network of alumni connections as an instrument in a two step approach: in the first step the cosponsorship network is "estimated" using the alumni network; in the second step the residuals obtained in the first step are used to control for endogeneity in a linear-in-means model of legislative effectiveness where the lawmakers' connections are proxied by the cosponsorhip network. As we discuss in greater detail in Section 5.2, a limitation of this approach is that, in the absence of a model of network formation, there is no obvious way to "estimate" a matrix of observed cosponsorship links with a matrix of alumni connections in the first step. In Battaglini et al. [2020b], the estimation is limited to an estimation of the direct binary links in the cosponsorship network as a linear functions of the binary link in the alumni network and other observables. Even if we accept the linearity of the homophily model in the first step, the simple pairwise estimation ignores important structural network characteristics of the social network. Our model of endogenous network formation incorporates these second order effects because each link in equilibrium reflects the equilibrium levels of effectiveness (which in turn depends on the entire network topology). Indeed, we show that our approach yields a significantly better fit to the data than this approach or other variants.

Three other papers in the networks literature deserve special mention. The first two are Konig [2016] and Boucher [2020], who also use an ABC approach to estimate a model of strategic network formation. In contrast to our approach, these papers estimate the probability distribution over networks under the assumption that a network realization is observed. The sequential models of social link formation underlying these works do not have a tractable analytical characterization of the equilibria, thus they do not give sufficient statistics for the approximation of the likelihood function. Konig [2016] and Boucher [2020] suggest alternative sets of summary statistics which allow them to estimate the knowledge spillovers using data on patents and scientific coauthorships among economists and physicists in the first paper; and homophily in a network of high school friends in the second. The second related paper is Canen et al. [2020], who adapts a model by Cabrales et al. [2011] to estimate social efforts in the U.S. Congress. In contrast to our paper, the authors assume that legislators cannot target their socialization efforts to other specific legislators, rather they exert a generic non-directed level of effort in socializing with all other legislators (Cabrales et al. [2011] call it a model without "earmarked socialization").[13] The empirical analysis, moreover, is based only on the total number of bills cosponsored by a legislator, an empirical proxy for social effort; and on a one-dimensional index of roll call data and floor speeches, an empirical proxy for legislative effort.

[13] A member of a legislator's party may benefit differently from the social effort, but this is not a choice of the legislator, it instead depends on an exogenous parameter modulating "partisanship."

2 Model

In Section 2.1, we describe the general model of endogenous network formation that will be used in the network estimation in Section 5. To keep the analysis concrete, we present the setup as a model of social connections between legislators in the U.S. Congress. This will indeed be the main application developed in Sections 5. In Section 2.2, we discuss other environments in which social networks are endogenous and unobserved to outsiders and explain how the model can be adapted to study these different environments.

2.1 Setup

Consider a Congress comprised of n legislators, where $\mathcal{N} = \{1, ..., n\}$ is the set of legislators. Each legislator has a pet legislative project that he or she wants to implement. The goal of each legislator is to maximize their legislative effectiveness, measured by the probability of implementing the project.[14] We assume that legislator i's effectiveness E_i is an increasing function of the effort directly exerted by i and the legislative effectiveness of all the legislators with whom i is socially connected. Legislator i's PAC, for example, may have contributed to the reelection campaign of legislator j, so j may use his or her "weight" to help i. Specifically, we assume the following "production function" for legislative effectiveness:

$$E_i = \rho \cdot (s_i)^{\alpha} (l_i)^{1-\alpha} + \varepsilon_i \tag{1}$$

The Cobb-Douglas function in (1) captures the effects of legislator i's level of "social connectedness" s_i and effort l_i. We assume that i's social connectedness is

$$s_i = \sum\nolimits_{j \in \mathcal{N}} g_{i,j} E_j, \tag{2}$$

where $g_{i,j}$ is a measurement of the social link between i and j. The idea behind (2) is that i's effectiveness is increasing in the effectiveness of his/her social connections within the legislature. Because of this, the effect of j on i is weighted in (2) by the degree of social connection of i to j. The second term, ε_i, is a factor idiosyncratic to i that contributes to i's efficacy independently from his/her connections or effort. We assume this factor is observed by the lawmakers but not by an econometrician studying the game. In the analysis below, we assume $g_{i,i} = 0$, $g_{i,j} \in [0, \overline{g}]$ with $\overline{g} > 0$, $\varepsilon_i \in [\underline{\varepsilon}, \overline{\varepsilon}]$ with $\underline{\varepsilon} > 0$, $\overline{\varepsilon} \in (0, 1)$, and $l_i \in \left[0, \overline{l}\right]$ with $\overline{l} > 0$. Moreover, below we will maintain the following assumption that guarantees $E_i \in [0, 1)$:

Assumption 1. $\rho \cdot \overline{g}^{\alpha} \cdot \overline{l}^{1-\alpha} + \overline{\varepsilon} < 1$.

[14]The idea that legislators have independent projects that they pursue to seek reelection is at the core of the theory of distributive politics (see Fiorina [1978] and Weingast [1979], among others).

These assumptions on the parameters and functional form are only made for convenience. In Section 6, we discuss how more general functional forms would affect the analysis.

In this model, legislators' effort levels $l = \{l_1, ..., l_n\}$, legislative effectiveness $E = \{E_1, ..., E_n\}$ and the social matrix $G = (g_{i,j})_{i,j \in \mathcal{N}}$ are all endogenous variables. These variables are determined in a two-stage game. At $t = 2$, the legislators choose their costly efforts l_i, taking the social links G as given. The cost of effort is assumed to be represented by a linear function $L_i(l_i) = c \cdot l_i$, where c is a cost parameter.

At $t = 1$, legislators befriend other legislators in order to increase their legislative effectiveness. At this stage, the legislators simultaneously choose the social links $g_{i,j}$. Specifically, we assume that at $t = 1$, legislator i decides with which other legislator $j \in \mathcal{N} \setminus i$ s/he wishes to establish a link $g_{i,j}$. A link $g_{i,j}$ depends on i's effort but is only established if j approves it. If a social link from i is approved by j, the cost of establishing it with intensity $g_{i,j}$ is given by:

$$C(g_{i,j}, \theta_{i,j}) = \frac{\lambda}{(1+\lambda)} \left(\frac{g_{i,j}}{\theta_{i,j}} \right)^{1+\frac{1}{\lambda}}, \tag{3}$$

where $\theta_{i,j}$ is a variable that captures the degree to which the types of i and j are socially "compatible:" the more i and j are socially compatible, the lower is the cost for i to establish a link with intensity $g_{i,j}$ with j. This cost may be interpreted as, for example, the cost of the time spent socializing with j or the time that i's staff needs to spend with j's staff in order to coordinate actions, or the cost of campaign contributions from i's PAC to j's PAC. The variable $\theta_{i,j}$ is taken as exogenous in the theoretical analysis and it may comprise a number of factors: whether i and j are elected in the same state, whether they have the same party affiliation, gender, or educational background (for example if they attended the same educational institutions). In practice, we assume that the matrix $\Theta = (\theta_{i,j})_{i,j}$ is symmetric and that for each legislator i there is a set $\mathcal{M}_i$ of other legislators such that $\theta_{i,j} > 0$ for $j \in \mathcal{M}_i$ and zero otherwise. This implies that legislator i is compatible with at most a subset $\mathcal{M}_i$ with cardinality $m_i = |\mathcal{M}_i|$ of other legislators. We denote $\overline{m} = \max_i m_i$ as the maximal cardinality of the subsets of friends. The variables $\theta_{i,j}$ and $\mathcal{M}_i$ will be discussed in greater detail in Section 4.1, when we develop the empirical analysis of the model. While Θ and $\mathcal{M}_i$ are exogenous in the theoretical model, the variables that determine them will be estimated in the empirical analysis.

In the socialization process described above, the ability of i to establish a link with j depends only on i's effort and i and j's types, not on j's effort. Naturally, it may be that j's effort plays a role too. In Section 6.1 we extend the model to allow $g_{i,j}$ to be a function of both i and j's effort, along with their types. The analysis is more complicated, but it is not qualitatively different.

The following assumption guarantees that we will not have a corner solution in which a legislator chooses $l_i = \bar{l}$ for some $i \in \mathcal{N}$.[15]

[15] A formal proof of this fact is provided in the proof of Proposition 1.

Assumption 2. $\bar{l} > ((1-\alpha)\,\rho/c)^{1/\alpha}$

Note that a simple condition that guarantees both Assumption 1 and Assumption 2 are satisfied is that the parameter controlling the social spillovers ρ is sufficiently small.

The type ω_i of a legislator i is defined by all of the variables describing his/her preferences and social connections, so $\omega_i = (\varepsilon_i, (\theta_{i,k})_{k\in\mathcal{N}}, \mathcal{M}_i)$. We denote with Ω the space of types with typical element $\omega \in \Omega$. A pure strategy for a legislator is described by a socialization strategy $g : \Omega \to [0, \overline{g}]^{n-1}$, mapping the legislator's type to a vector of intensities $g_i = \{g_{i,j}\}_{j\neq i}$ for each of the $n-1$ other legislators; and an effort strategy $l : \Omega\times G\to \left[0, \bar{l}\right]$, mapping the social network and i's type to an effort level.

An important feature of the process of network formation presented above is that we do not assume that the choices of social connections occurs in a vacuum. We assume that in the background there is an infrastructure of preexisting relationships that may affect the cost of establishing (or maintaining) social links and which are potentially observable. In the model, this infrastructure enters through the variables $(\theta_{i,j})_{i,j}$ that affect the costs of forming a link between i and j. In the empirical implementation presented below, we will allow these variables to depend on observable social networks such as alumni connections, or a variety of other measures of social and geographic distance. We take this social infrastructure as exogenous, therefore, but we allow the agents to build on it. We think it is natural to assume that at any point in time an agent is endowed with a given framework of social relations upon which s/he can build his/her relevant social network. There is indeed a dialectic relationship in our model between the relevant *endogenous network* that affects the social outcomes of interest, and a variety of potentially relevant *exogenous networks* that may be relevant for the establishment of social connections. This distinction between the endogenous and exogenous social networks is useful because we are often directly interested in the shape of the underlying social network, not just in establishing the existence of social spillovers. Observable measures of social distance, however, are unlikely to coincide with the social network relevant for the outcome of interest. Still, they might be important ingredients in its endogenous determination and therefore could be useful in its estimation. The distinction is also useful as a building block of a dynamic theory of network formation, in which the endogenous network at t enters as a state variable (exogenous network) in the formation of the relevant network at $t+1$. We will discuss in greater detail such an extension in Section 6 and in Section A.1 of the online appendix.

2.2 Alternative interpretations

As mentioned above, the model is amenable to alternative interpretations. Here, we discuss some of these possible alternative applications. Other examples will be discussed in Section 3.4 after we present the results.

2.2.1 Adolescent behavior

A large recent literature has studied how adolescent behavior depends on social connections. The focus has been on whether adolescents' levels of educational achievements, risky behavior and criminal behavior depends on peers. Adolescent's social connections are clearly endogeneous and not directly observed, at least by outsiders. The literature bypassed this problem using a variety of proxies for social connectedness.[16] There are, however, three issues with these approaches. First, they may be imprecise and leave out important connections. This is especially true if links are defined at the school or classroom level because adolescents are likely to cultivate relationships outside the classroom that may significantly alter the topology of their social networks. It may also be true for self reported connections because adolescents may be reluctant to disclose all of their relationships.[17] Second, detailed self reported friendship relations are available only for specific data sets and for the specific periods in which the surveys were conducted. Finally, we might have multiple sources of information providing different measures of social connectedness (such as class membership, gender, after school activities, and neighborhood), but we might not know how to combine them.

The framework presented in Section 2.1 can be immediately applied to these settings. For example, consider Calvo-Armengol et al. [2008], who present a model in which educational achievements y_i are shown to depend on the levels of educational achievements of connected peers. In their microfoundation, the following relationship holds:[18]

$$y_i = \mu g_i + \theta_i(x) + \phi \sum_{j \in N}^{n} g_{i,j} y_j, \tag{4}$$

where $g_{i,j} = 1$ if i and j are direct friends and $g_{i,j} = 0$, otherwise; $\theta_i(x)$ is a function of observable characteristics x and ϕ is a parameter; and $g_i = \sum_j g_{i,j}$. If we interpret E_i as the level of education achievement y_i and we allow the $g_{i,j}$ to take continuous values, condition (4) can be interpreted as a special case of the production function (1) in which we do not allow for endogenous effort by the agents (i.e. where $\alpha = 1$) and the idiosyncratic term ε_i has mean $\mu g_i + \theta_i(x)$.[19]

In the context of Calvo-Armengol et al. [2008], however, it is natural to allow students' links to be endogenous and to depend on the appeal in equilibrium of the potential partners. Relying

[16]For example, Hanushek et al. [2003] and Angrist and Lang [2004] used school membership; Kang [2007] and Boucher et al. [2014], among others, relied on classroom membership; Calvo-Armengol et al. [2009] and Patacchini and Zenou [2012] exploited self reported friendship relations in the National Longitudinal Survey of Adolescent Health (Add Health).

[17]About 30 percent of adolescents in the Add Health survey report having no friends. Moreover, a large fraction of reported friendships are not corresponded by the cited partner.

[18]The condition corresponds to equation (28) in Calvo-Armengol et al. [2008].

[19]Indeed, in the empirical implementation in Section 5, we will assume $\varepsilon_{i,t} = \mathbf{X}_{i,t}\beta + \zeta_t + \epsilon_{i,t}$, where β is a vector of coefficients, ζ_t denotes time fixed effects and the $\epsilon_{i,t}$s are random i.i.d. variables with zero mean. Even if we allow $\alpha \leq 1$ and we endogenize l_i, we obtain a characterization for E_i in the second stage, so for a given G, that is remarkably similar to (4) (see, for example, equation (6)).

on the NCE defined above, this goal can be achieved by augmenting their model with a first stage in which $G = [g_{i,j}]$ is selected as in Section 2.1.[20] As mentioned in the previous section, an important advantage of our approach is that it allows us to merge information from different observable adjacency matrices (such as class attendance, self reported friendship relations, etc) by incorporating this information in the distribution of the $\theta_{i,j}s$. This way we can potentially obtain a more accurate estimate of social connections.

2.2.2 Financial networks

With the term financial networks, we mean the web of business links that banks and other financial operators (such as hedge funds, insurance companies, venture capital funds, and others) establish among themselves: credit lines, derivatives trades, etc. Financial networks have long been recognized as important factors in financial crises, for interbank liquidity, and the diffusion of investment choices (see Allen and Babus [2009] for a survey of this literature). Financial networks are not only endogenous, but typically unobservable too. For example, many papers studying financial links between banks rely on credit exposures in the interbank market (see for example Drehmann and Tarashev [2011], Denbee et al. [2020]). Credit exposures, however, are not perfectly observed and results are often sensitive to how they are measured (see Upper [2006]). Two issues make the observation of interbank connections problematic. First, exposures are generally recovered from the banks' balance sheets which only provide information on the aggregate exposure of a bank.[21] Second, an important part of interbank links derives from off-balance sheet positions and these positions are not observable. As shown by Allen and Gale [2000], among others, contagion is very sensitive to the shape of the financial network. For regulations on financial stability, monetary policy, and policy intervention more generally, it is therefore important to incorporate an estimate of the real underlying links between the financial operators as unbiased and complete as possible.

To see how to apply our model to these settings, consider Denbee et al. [2020] who proposed the following model to study how interbank liquidity depends on the overall network of financial links. Let L_i be the level of liquidity of bank i. The authors postulate the following relationship:

$$L_{i,t} = \alpha_t + \alpha_i + \rho \sum_{j=1}^{n} g_{i,j} X_{j,t} + \rho \sum_{j=1}^{n} g_{i,j} L_{j,t} + \epsilon_{i,t} \tag{5}$$

where $X_{i,t}$ is a vector of observables, $g_{i,j}$ is the link between bank i and bank j, and $\alpha_{i,t}$ and α_t are individual and time fixed effects.[22] Once again, this model can be seen as a special version of

[20] As we will see in Section 3, when this is done we obtain a generalized version of (4) where the relationship between y_i and the y_j of another player depends on the elasticity of the link $g_{i,j}$ with respect to the effectiveness E_j, as measured by λ. See equation (17).

[21] Bilateral credit positions vis-a-vis the other banks are generally reconstructed using ad hoc algorithms such as the maximum entropy method (which is based on the idea that banks spread their position uniformly across other banks). Among others, the maximum entropy algorithm has been used by Drehman and Tarashev [2011] and Peltonen et al. [2015]. See Mistrulli [2007] for a survey and discussion of these approaches.

[22] See Equation (20) in Denbee et al. [2020]. In their model, liquidity is denoted l_i. We use the capital letter

(17) in which the elasticity of link formation λ is pinned at 0. In Denbee et al. [2020] the financial network between banks is taken as exogenous and approximated using interbank borrowing data reconstructed using a specific algorithm from Furfine [2003]. Equation (5) can be seen as the outcome of the second stage of our game (in which liquidity is determined given the network G). We can now, however, extend (5) as we do in the first step of our model by endogenizing the network. In this case, bank i selects $g_i = (g_{i,1}, ..., g_{i,n})$ to maximize its liquidity $L_{i,t}$. Data on interbank relationships (such as the Furfine [1999] algorithm) along with other sources of information can now be used as inputs in $\theta_{i,j}$ to estimate the latent G.

2.2.3 Production networks

There is an established literature studying input-output linkages among firms and their impact on the propagation of shocks (see Carvalho and Tahbaz-Salehi [2019]). In empirical analyses of the U.S., this literature generally relies on the input-output accounts data complied by the Bureau of Economic Analysis (BEA) (see Carvalho [2014]). While this dataset breaks down the data in a relatively granular way, it aggregates the values at the industry level. Data at the firm level is available only for a few countries (such as Japan and Belgium).[23] Comparative analyses of the data at the firm and industry levels have highlighted features of firm-level networks that are not apparent in aggregated data, suggesting that it would be valuable, for countries where firm-level data is not available, to estimate the determinants of intra-industry heterogeneity in firm connections using outcomes as in our approach.

While this literature has recognized the importance of homogenizing the input-output production matrices (see for instance Oberfield [2018]), the research on this front has been primarily theoretical. Empirical applications have been based on models calibrated under the assumption that input-output links are observable. There has been no attempt to recover the input-output matrix under the assumption that it is unobservable. The approach developed in our paper may contribute to this literature by providing a set of tools to recover the determinants of firms' heterogeneity. The model would use observable outcomes (such as productivity per worker), available firm-level data (location, history, size for example) and industry-level data to provide an estimate of link connections at the level of the firm. The model described in Section 2.1 can be applied to these frameworks directly if we interpret E_i as the productivity of a firm i. More generally, even allowing for alternative microfoundations for the productive externalities, the concept of the Competitive Network Equilibrium may prove useful to endogenize the links between firms.

here to distinguish this variable from the effort levels $l_i s$ defined in our model.

[23]For Japan, it is available as a large private dataset from a credit reporting firm named *Tokyo Shoko Research* (TSR) that collects data for the universe of firms with more that five employees. For countries with the value-added tax (VAT), such as Belgium, data on input-output relationships can be obtained from tax reports.

2.2.4 Board connections

A recent literature studies whether social connections of CEOs, corporate directors, and portfolio managers affect the performance of the company where they work, and/or how they matter for their careers.[24] It is clear that social connections between managers (those that matter for transmission of information, or for the sharing of ideas, ethical standards, and habits) can only be roughly approximated by the observable measures used in the literature and that if anything these measures should be combined to obtain a more accurate measure rather than used independently. It is also clear that social connections among boards of directors and CEOs are endogenous. The very fact that so many different measures of social connectivity have been used to measure links between businessmen shows that they are not directly observable. The model presented in Section 2 can be immediately applied to study these situations if we interpret the levels of effectiveness E_i as the effectiveness of a CEO (which we might be proportional to his/her compensation, likelihood of turnover, and/or performance of the firm). Managers endogenously establish links among themselves in a variety of ways, for example by inviting each other to join their boards of directors (as Steve Jobs did when he invited in 2006, and subsequently dis-invited in 2009 Eric Schmidt of Google to join Apple's board). In this context, a variety of observable factors affecting social connectivity (participation on the same boards, alumni connections, etc) can be used for recovering the true latent social network.

3 Equilibrium analysis

3.1 Network competitive equilibrium

The game described in the previous section has a simple structure that allows us to solve it by backward induction. At $t = 2$, the legislators choose effort levels while taking the social network as given; at $t = 1$, legislators choose their social links. As it is often the case in games with network externalities, however, the analysis is complicated by the fact that each action has both a straightforward direct effect and a set of indirect effects. For example, consider the choice at $t = 1$, when legislator i chooses the link to j, $g_{i,j}$. Here a change in $g_{i,j}$ has a direct effect on E_i described by (1), but it may also have a complex set of indirect effects. The change in E_i given G changes all other E_ls of ls who are connected to j, and these changes may affect E_j if j is connected to them, directly or indirectly.

To understand our approach, it is useful to note that these complications are not dissimilar to the complications that arise when studying a general equilibrium in an exchange economy in

[24] Hochberg et al. [2007] has studied whether the relationship established by venture capitalists with other operators (VCs, lawyers and other consultants) in the organization of syndicated loans and investments affect the performance of their respective funds. Cohen et al. [2008] have used alumni connections of portfolio managers and corporate board members to study if social links allow portfolio managers to acquire useful information for their investments.

which a change in an agent's demand has a direct obvious effect on an agent's utility and an indirect effect on equilibrium prices. The solution in general equilibrium analysis is to assume that agents are "price takers:" agents solve their optimization program taking prices as given; prices, however, must clear the market in equilibrium. Such analysis is motivated by the fact that, in many exchange economies, each agent has only a marginal impact on equilibrium prices, thus allowing us to ignore the indirect effects.

The same approach seems appropriate for the study of network games with many players such as ours, in which the incentives to establish a link to a node depends only on some measure of centrality of the node that is a function of the aggregate behavior in the network. In these environments it is plausible to assume, as in a competitive equilibrium, that the players are "price takers" with respect to these measures of centrality. In our case, the centrality measures are the legislators' levels of effectiveness. We can, therefore, introduce a *Network Competitive Equilibrium* as follows:

Definition 1. *Legislators' effort levels* $l = \{l_1, ..., l_n\}$*, legislative effectiveness* $\mathbf{E} = \{E_1, ..., E_n\}$ *and the social matrix* $G = (g_{i,j})_{i,j \in \mathcal{N}}$ *are a Network Competitive Equilibrium (NCE) if:*

- *network connections* $\mathbf{g^i} = (g_{i,1}, .., g_{i,n})$ *are optimal for* i *at t=1 given* $\mathbf{E}$*;*
- *effort levels* l_i *are optimal for legislator* i *at* $t = 2$ *given* $\mathbf{E}$ *and* $G = \left(\mathbf{g^i}\right)_{i \in \mathcal{N}}$ *;*
- *the vector of efficacy levels* $\mathbf{E}$ *satisfy the production function (1) given* $\mathbf{l}$ *and* G*.*

The first two requirements in the definition correspond to the requirement in a competitive equilibrium that agents optimize given "prices," where the other legislators' levels of effectiveness correspond to prices. The last requirement corresponds to the market clearing condition: here we impose that the equilibrium expected levels of effectiveness are consistent with each other. It should therefore be stressed that while the legislators choose G taking $\mathbf{E}$ as given, $\mathbf{E}$ is endogenous in the same way as prices are endogenous in a competitive equilibrium. We will explore some properties of this equilibrium definition below.

While conceived for the game described in Section 2, the equilibrium concept in Definition 1 has general applicability in games with endogenous network formation. It applies to any environment in which the benefit of establishing a link with a node depends on a measure of centrality of the node (in our case the effectiveness, which as we will see corresponds to a weighted Bonacich centrality). The novelty of the approach is to assume that the players take this synthetic measure of the importance of a node as given when choosing the links.

Before turning to the equilibrium characterization, it is useful to compare our approach with the other approaches that have been adopted to model network formation for empirical analysis. As mentioned above, the key aspect of our NCE is the simplification of the strategic interaction implied by the assumption that agents take the other players' effectiveness as given when choosing their social connections. Underlying this approach is the implicit assumption that agents are "small" and thus have (or perceive) only a marginal effect on the effectiveness of other players. This assumption does not seem too demanding in large and complex networks such as congressional networks. Consider, for example, Figure A.1 in the online appendix, plotting our estimated network and other observed networks often used to study social networks in Congress (such as the cosponsorship network).[25] It is apparent that no nodes (or very few) are in a position to exert a dominant effect on social interactions; and, more importantly, that if we change an individual agent's connections, we would not change the overall network very much.

The other two approaches to model endogenous networks used in the literature–stochastic best response dynamics and pairwise stability–also adopt very significant simplifications of the strategic interactions between agents. In the first approach, based on stochastic best response dynamics, agents are myopic and form random links sequentially. This approach, moreover, does not lead to an analytical characterization of an equilibrium (except under simplifying assumptions in specific examples). The equilibrium stationary distribution is obtained by simulating the underlying Markov process. The underlying Markov process typically induces a unique stationary distribution of networks, but this distribution depends on the exogenous distribution of the random shocks that affect the agents' preferences. The second approach focuses on networks that satisfy pairwise stability, a cooperative solution concept. This approach typically leaves a large set of equilibrium predictions and thus only partial identification under strong sparsity conditions on the network.[26]

While the NCE can conceptually be applied to any network, its applicability is probably not appropriate for simple environments with few links, where it is likely the case that players are well aware of indirect effects (just as the idea of a competitive equilibrium can conceptually be applied even to a one agent Robinson Crusoe economy, but probably it should not). Whether the simplification in the NCE is an acceptable compromise is ultimately an empirical question that has to do with the ability of the model to fit the data better than alternative approaches. We will argue in Section 5 that this is the case for our congressional network.[27]

In the next two subsections we characterize the NCE of our game, starting from the choice of effort at $t = 2$.

[25]For a definition of the cosponsorship network, see footnote 13. The committee network is the network in which two lawmakers are connected if they serve on the same committee

[26]For empirical studies following this approach, see Miyauchi [2016], De Paula et al. [2018], and Sheng [2018], among others.

[27]The NCE, however, can be extended to study environments in which there is an elite of non-atomistic players who influence the behavior of other players and are strategic about it; and there is a set of atomistic players who are "price takers" in the sense described above. We will describe this extension in Section 6.

3.2 The choice of effort at $t = 2$

Substituting the optimal level of effort $l_i(G, \varepsilon)$ at $t = 2$ into (1), we obtain that the equilibrium levels of legislative effectiveness for a type $i \in N$ are given by:

$$E_i = \rho \left(\frac{(1-\alpha)\,\rho}{c} \right)^{\frac{1-\alpha}{\alpha}} \cdot \sum\nolimits_{j=1}^{n} g_{i,j} E_j + \varepsilon_i. \tag{6}$$

These equations can be expressed in matrix form as:

$$[I - \delta \cdot G] \cdot \mathbf{E} = \varepsilon \tag{7}$$

where $\delta = \rho\left((1-\alpha)\,\rho/c\right)^{\frac{1-\alpha}{\alpha}}$ and ε is the vector $(\varepsilon_i)_{i \in \mathcal{N}}$. If we had an exogenous G, condition (4) would have a straightforward interpretation: assuming the invertibility of the matrix on the left hand side of (7), a legislator's effectiveness coincides with his weighted Bonacich centrality in G, with weights given by the natural "effectiveness of each legislator" ε_i and the discount factor δ.[28] However, this interpretation of E as "weighted Bonacichs" is no longer correct because G is endogenous.

3.3 The formation of the network at $t = 1$

At $t = 1$, the legislators choose their social links to maximize the expected utility at stage $t = 2$, net of the cost of establishing the links. The expected continuation utility at $t = 1$ of a type i (i.e. $E_i(G, \varepsilon) - c l_i(G, \varepsilon)$ for i) is easily determined as:

$$U^i(G, \varepsilon) = \alpha\delta \sum\nolimits_{j=1}^{n} g_{i,j} E_j(G, \varepsilon) + \varepsilon_i. \tag{8}$$

Legislator i will choose the links $\mathbf{g^i} = (g_{i,1}, ..., g_{i,n})$ that maximize (8) with the constraint that if $g_{i,j} > 0$, then the link is not vetoed by j. It is, however, easy to see that no legislator j would find it optimal to veto a link from i. The establishment of a link $g_{i,j}$ increases the effectiveness of i and of any other legislator who has a direct or indirect link to i: so if j does not have a direct or an indirect link that points to i, then j is indifferent; if j has a direct or indirect link to i, then j strictly prefers that i establishes a link with him/her.[29] It follows that legislator i chooses his links by solving:

$$\max_{\mathbf{g^i}} \left\{ \sum\nolimits_{j=1}^{n} \left[\alpha\delta \cdot g_{i,j} E_j(G, \varepsilon) - \frac{\lambda}{(1+\lambda)} \left(\frac{g_{i,j}}{\theta_{i,j}} \right)^{1+\frac{1}{\lambda}} \right] \right\}. \tag{9}$$

Combining the solution of (9) with (7), we have:

[28] The standard definition of the Bonacich centrality with discount factor ν is $\mathbf{E} = [I - \nu \cdot G]^{-1}\,\mathbf{1}$. The weighted Bonacich with weights $\mathbf{A}$ is defined as $\mathbf{E} = [I - \nu \cdot G]^{-1}\,\mathbf{A}$. See Ballester et al. [2006].

[29] In Section 6.1, we will extend the analysis to consider the case in which $g_{i,j}$ depends on both i's and j's investments.

Proposition 1. *A Network Competitive Equilibrium (NCE) exists and it is characterized by a vector E^* and a matrix G^* that solve the system:*

$$E_i^* = \delta \cdot \sum_{l \in \mathcal{N}} \left(g_{i,l}^* E_l^* \right) + \varepsilon_i \tag{10}$$

$$\text{and } g_{i,j}^* \leq (\theta_{i,j})^{1+\lambda} \left(\alpha \delta E_j^* \right)^{\lambda} \ \ (= \text{ for } g_{i,j}^* \leq \overline{g}) \tag{11}$$

for any $i, j \in N$.

In an interior solution, i.e. when $g_{i,j}^* \leq \overline{g}$ for all i, j, the two conditions collapse to the system of n equations and n variables:

$$E_i^* = \alpha^{\lambda} (\delta)^{1+\lambda} \sum_{l \in \mathcal{N}} (\theta_{i,l} E_l^*)^{1+\lambda} + \varepsilon_i \tag{12}$$

The legislators' effectivenesses are no longer representable by a linear system of equations as in the familiar "Bonacich" representation of (7). The intuition for this phenomenon is simple. When the network is exogenous, E_i is a linear function of E_j, with a factor of proportionality given by $g_{i,j}$. When $g_{i,j}$ is endogenous, however, i finds it optimal to choose $g_{i,j}$ that is proportional to $(E_j)^{\lambda}$. The true link between E_i and E_j, therefore, is no longer linear: E_i will be a function of $(E_j)^{1+\lambda}$.

To interpret (10) and (12), it is useful to note that the elasticity of a link $g_{i,j}$ with respect to the effectiveness of the associated target legislator j is $\epsilon_{g_{i,j},E_j} = \lambda$.[30] As $\lambda \to 0$, the endogenous links become completely inelastic with respect to effectiveness, and indeed we have $g_{i,j} \to \theta_{i,j}$. In this case, we are back to the standard Bonacich representation of effectiveness, assuming that $[I - \delta \cdot \Theta]$ is invertible:

Example 1. *As $\lambda \to 0$, there is a unique equilibrium in which effectiveness coincides with the Bonacich centralities: $\mathbf{E} = [I - \delta \cdot \Theta]^{-1} \cdot \varepsilon$, where Θ is the $n \times n$ matrix with generic term $\theta_{i,j}$.*

When $\lambda > 0$, instead, changes in the equilibrium effectiveness imply changes in the links. The higher is λ, the more links polarize around the most effective legislators. To see the implications of a positive λ, let $\lambda \to \infty$. In this case, from the first order condition, we obtain that $g_{i,j} = \overline{g}$ if $\alpha \delta E_j \theta_{i,j} - 1 \geq 0$ and zero otherwise. Consider an environment in which legislators are symmetric (so ε_i is the same for all i, say at $\varepsilon_i = \varepsilon$) and located in a ring such that $\theta_{i,j} = 1$ for $j \in \{i-1, i+1\}$ and zero otherwise. We say that there is no connectivity if $g_{i,j} = 0$ for all $i, j \in \mathcal{N}$ and that there is full connectivity if $g_{i,j} = \overline{g}$ whenever $\theta_{i,j} > 0$. We have:[31]

Example 2. *As $\lambda \to \infty$, in a ring there is a unique pure strategy equilibrium with no connectivity if $\varepsilon < 1/(\alpha\delta)$; a unique pure strategy equilibrium with full connectivity if $\varepsilon \geq (1 - 2\delta\overline{g})/(\alpha\delta)$; and both equilibria coexist if $\varepsilon \in [(1 - 2\delta\overline{g})/(\alpha\delta), 1/(\alpha\delta)]$.*

[30] The elasticity is defined as $\epsilon_{g_{i,j},E_j} = (\partial g_{i,j} / \partial E_j) \cdot (E_j / g_{i,j})$.

[31] The proof of the statement in this example is presented in Section A.2 in the online appendix.

The network structure described in Example 1, where links are inelastic to the level of effectiveness of the legislators (i.e. $\lambda = 0$) and thus exogenous, is very different than the structure in Example 2. While the network formation decision is continuous in the first case, it shows an "explosive behavior" in the second (as a function of ε, at least).

In the following, we will assume that we do not have corner solutions with $g_{i,j} = \overline{g}$ and so the equilibrium network is characterized by (12). This property is implied by the following assumption on the fundamentals:

Assumption 3. $\overline{g} > (\alpha\delta)^{\lambda}\,\overline{\theta}^{1+\lambda}$.

Under Assumptions 1-3, it is easy to state a sufficient condition for the existence of a unique equilibrium for δ sufficiently small. Define $\overline{\theta} = \max_{i,j\in\mathcal{N}} \theta_{i,j}$. We have:

Proposition 2. *For $\delta < \frac{1}{\overline{\theta}}\left[1/\left((1+\lambda)\,\alpha^{\lambda}\overline{m}\right)\right]^{1/(1+\lambda)}$, there is a unique equilibrium $G^* = (g^*_{i,j})_{i,j\in\mathcal{N}}$, $E^* = (E^*_i)_{i\in\mathcal{N}}$.*

It should be noted that the condition in Proposition 2 is only sufficient, not necessary. For example, the symmetric example presented above with arbitrarily large λ obviously violates this condition but still admits a unique equilibrium under these specified conditions.

3.4 What goes wrong when ignoring unobservability and endogeneity?

Before turning to the empirical specification of the model, we discuss three problems that may emerge when ignoring the fact that the true network is unobserved and endogenous. Example 1 illustrates what we call the *extensive margin* problem, which emerges when only part of the true network is observable. Example 2 illustrates what we call the *intensive margin* problem, which emerges when even if we observe the presence of social links, we do not observe their intensity or strength. Finally, in the third example, we discuss the *multiplicity of sources problem*, which emerges when we have multiple observable social matrices that can be used as proxies of the true network but we do not know which is the relevant one (or, if more than one is relevant, which weights to use when combining them).

Example 1: the extensive margin

Assume we have $2m$ students, denominated $b_{l,B}$, $b_{l,G}$ for $l = 1, ..., m$: where l stands for classroom $l = 1, ..., m$; G stands for good type and B for bad type (a kid is identified by $\tau = (l, t)$ where $l = 1, .., m$ is the class and $t = G, B$ is the type; the set of kids is T). The students' educational achievements $E = (E_\tau)_{\tau\in T}$ follow the production function: $E_\tau = a_\tau + \rho\sum_{\tau'\in T} g_{\tau,\tau'} E_{\tau'}$, meaning that they depend on the intrinsic ability a_τ, and on links to other students.[32] Links between classmates are observed and equal to 1: that is, $g_{\tau,\tau'} = 1$ for any $\tau = (i, t)$ and $\tau' = (j, t')$ who

[32] This production function corresponds to (1) where for simplicity we set $\alpha = 1$.

are in the same class (so $i = j$) for any $t, t' \in \{G, B\}$. Suppose now that bad students have a hard time socializing with good students in different classes; and good students have a hard time socializing with bad students in different classes: formally, $\theta_{\{l,B\},\{k,G\}} = \theta_{\{i,G\},\{j,B\}} = 0$ for all l, k and i, j such that $l \neq k$ and $i \neq j$. Instead, $\theta_{\{i,B\},\{j,B\}} = \theta_{\{l,G\},\{k,G\}} = \kappa > 0$ for all couples i, k and l, k such that $i \neq j$ and $l \neq k$. Socialization is therefore possible for students of the same type in different classes. When there is socialization between students in different classes, however, the links are unobserved. The social network, for a simple example with $m = 2$, looks as in Figure 1, where the solid line represents an observable and exogenous social link, while a dashed arrow represents a potential observable link.

We can use the characterization of Sections 3.2-3.3 to solve the model. The students' educational attainments solve the system:

$$\begin{aligned} E_{l,G} &= a_G + \rho[E_{l,B} + m\rho^{\lambda}(\kappa)^{1+\lambda}(E_{l,G})^{1+\lambda}] \\ E_{l,B} &= a_B + \rho[E_{l,G} + m\rho^{\lambda}(\kappa)^{1+\lambda}(E_{l,B})^{1+\lambda}], \end{aligned} \tag{13}$$

where we have simplified the model assuming $\alpha = 1$ so that $\delta = \rho$.[33] System (13) admits a unique solution $E^* = (E^*_{l,G}, E^*_{l,B})$ equal to $(0.2427, 0.1331)$ if we set in this example $\rho = 0.1$, $\kappa = 1$, $\lambda = 1$ and $m = 50$. Our goal is to retrieve the underlying fundamental parameters of the model using only the observable outcome E^*.

Following the existing literature discussed in Section 2, suppose first, as a benchmark, that we ignore that G types can make friends in classes different than theirs and instead we allow our analysis to rely on the only observable proxy for social connectedness: classroom membership. In this case, we assume $g_{\{i,G\},\{i,B\}} = 1$ for $i = 1, 2$ and $g_{\{j,k\},\{j,l\}} = 0$ otherwise. Assuming no noise here for simplicity, this gives us the system: $E_{l,G} = a_G + \rho E_{l,B}, E_{l,B} = a_B + \rho E_{l,G}$. If we estimate ρ by minimizing the quadratic error of the model, we have:

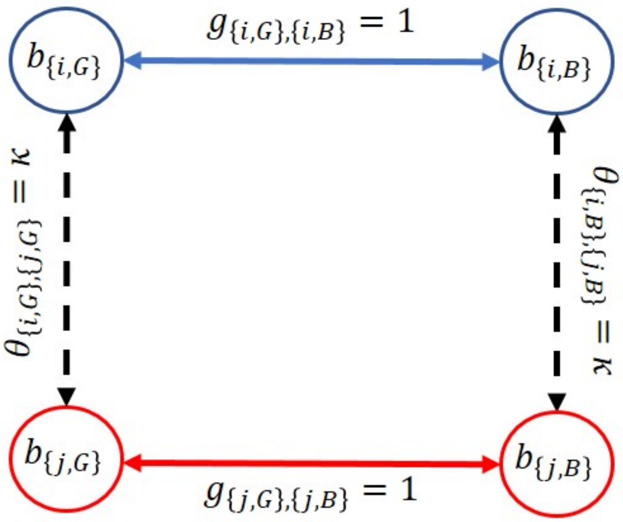
$$\widehat{\rho} \in argmax_{\rho} \left\{ K \cdot [(E^*_{l,G} - a_G - \rho E_{l,B})^2 + (E^*_{l,B} - a_B - \rho E_{l,G})^2] \right\} \tag{14}$$

It is easy to solve for the estimated value that, unsurprisingly, is significantly incorrect: $\widehat{\rho} = 0.179 > \rho = 0.1$. The problem here is one of the "extensive margin," in the sense that we are ignoring the link across classes, links that fundamentally change the topology of the network, and thus overestimating the marginal effects of the remaining links (but underestimating the total effect because there are fewer links).

Naturally, the key question is whether, when recognizing the endogeneity of the links, we can

[33]Because the link between classmates is exogenously given in this example, a student (i, t) endogeneously selects only the link with (j, t) for $j \neq i$ and $t = G, B$.

Figure 1: Example 1 - setup



Notes. The observable social network in Example 1 is illustrated by the solid lines. The unobservable network is illustrated by the dashed lines.

retrieve the fundamentals using only E^*. The problem is now three-dimensional:

$$\left(\hat{\rho}, \hat{\kappa}, \hat{\lambda}\right) \in argmin_{\rho,\kappa,\lambda} \sum_{l=1}^{m} \left[\begin{array}{c} \left(E_{l,G} - a_G - \rho\left(E_{l,B} + m\rho^{\lambda}(\kappa)^{1+\lambda}(E_{l,G})^{1+\lambda}\right)\right)^2 \\ + \left(E_{l,B} - a_B - \rho\left(E_{l,G} + m\rho^{\lambda}(\kappa)^{1+\lambda}(E_{l,B})^{1+\lambda}\right)\right)^2 \end{array} \right] \tag{15}$$

for a generic choice of $a_{l,G}, a_{l,B}$ for $i = 1, 2$. It is easy to solve this problem and verify that there is a unique solution that coincides with the true parameters. The right panel of Figure 2 illustrates this point by plotting the level curves of the objective function in (15) and highlighting the true parameters as the black dot (which indeed minimizes the function).[34] The left panel of Figure 2 illustrates the value of the objective function in (14), highlighting the fact that the true parameters do not minimize it. In real estimation problems, identification is more complex because of random errors. As we will show in Section 4.3, where we present a number of simulations, even in these cases the distribution of the parameters can be retrieved with accuracy.

Example 2: the intensive margin

A different type of problem emerges when the issue is not (or, at least, not only) omitted links. Suppose that the set of links is known, but the intensity of the links is heterogeneous across members of the group. The *relative* intensity of the links is of particular importance because it can bias the estimate of ρ if it is ignored. The endogenous links are non linear functions of the $\theta_{i,k}$, so failure of modeling endogeneity can lead to use a network that, although observable, is

[34]While we represent the level curves in a two dimensional space, the true vector of parameters minimizes (15) in the full three-dimensional space.

significantly different from the true network.

Let us imagine we have 4 players in a network with heterogeneous $\theta_{i,j}$ and ε_is, drawn from some distribution. We can obtain the equilibrium levels of effectiveness $E^* = (E_1^*, ..., E_4^*)$ by solving (10)-(11). We have:

$$\frac{g_{i,j}^*}{g_{k,j}^*} = \frac{(\theta_{i,j})^{1+\lambda}\left(\alpha\delta E_j^*\right)^{\lambda}}{(\theta_{k,j})^{1+\lambda}\left(\alpha\delta E_j^*\right)^{\lambda}} = \left(\frac{\theta_{i,j}}{\theta_{k,j}}\right)^{1+\lambda}$$

When the links are endogenous, the relative intensity of the links is therefore magnified relative to any observable measure we may use to approximate them. If $\lambda = 1$ and $\theta_{i,j}$ is double the size of $\theta_{i,k}$, then $g_{i,j}$ will be approximately four times the size of $g_{i,k}$. If $\lambda = 2$ and $\theta_{i,j}$ is still double the size of $\theta_{i,k}$, then $g_{i,j}$ will be approximately eight times larger than $g_{i,k}$.

As in the previous example, it can be verified that if we adopt the linear model $E_i = \rho\left[\sum_j \theta_{i,j} \cdot E_j^*\right] + \varepsilon_i$, the estimate of ρ obtained minimizing $\sum_i \left(E_i^* - \rho\left[\sum_j \theta_{i,j} \cdot E_j^*\right] - \varepsilon_i\right)^2$ would be mistaken. If instead we allow for the full model, we can precisely recover the entire vector of parameters using E^*. Note that it would be impossible to estimate the true network from E^* without a model of endogenous network formation: even in this very simple environment, the network comprises $n(n-1) = 12$ links and we have only four observations. The link intensities instead can be estimated using E^* because the model contains only two parameters (in this example).

Example 3: the multiplicity problem

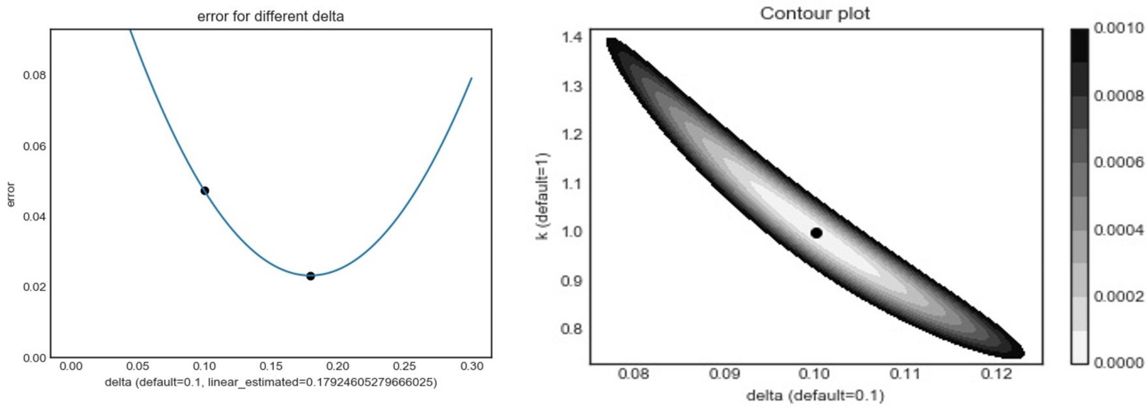
We conclude with an example to illustrate the benefits of combining alternative sources of information to analyze social interactions. Consider the model of Example 1 with four students $\left\{b_{\{1,G\}}, b_{\{1,B\}}, b_{\{2,G\}}, b_{\{2,B\}}\right\}$ in two classrooms. Now assume that the links between the G types across classes are indeed observable because, say, the two Gs attend the same chess club, and the two Bs are members of the same beer pong club. We then have the "classroom" network, along with the "after school" club network. Each student i has a best friend $F(i)$ who may be from the same class or from the after school club.[35] The question now is that we do not know whether the relevant observable link is the classroom, the club, or both.[36]

Suppose that the econometrician observes school classes and the memberships in the clubs but is unable to observe the relative importance of the clubs relative to the classroom. The econometrician may assume that the relevant network is only one of the two (say the school

[35] Formally, we have $\theta_{\{l,B\},\{k,G\}} = \theta_{\{i,G\},\{j,B\}} = 0$: but now we either have $\theta_{\{1,G\},\{2,G\}} = \theta_{\{1,B\},\{2,B\}} = 0$ and $\theta_{\{1,G\},\{1,B\}} = \theta_{\{2,G\},\{2,B\}} = 1$; or $\theta_{\{1,G\},\{2,G\}} = \theta_{\{1,B\},\{2,B\}} = 1$ and $\theta_{\{1,G\},\{1,B\}} = \theta_{\{2,G\},\{2,B\}} = 0$.

[36] Indeed, the problem may be much more severe because the students may attend a variety of clubs (karate, piano, etc). In total, even with just four students, we have six possible ways of partitioning the students in groups of two and thus six different possible social networks (the number of partitions in a group of four in groups of two is $4C2$, where nCm is the combination of n students in groups of m). If we allow for the union of only two partitions (considering, for example, the network $\left\{b_{\{1,G\}}, b_{\{1,B\}}, b_{\{2,G\}}\right\}$, union of $\left\{b_{\{1,G\}}, b_{\{1,B\}}\right\}$ and $\{b_{\{1,G\}}, b_{\{2,G\}}\}$), we arrive to 31 possible network configurations and we are not even considering the possibility of unpaired kids.

Figure 2: Example 1 - level curves of the nonlinear estimation



Notes. The left-hand side of the figure illustrates the value of the objective function in (14). The left-hand side represents the value of the objective function in (15) for $\lambda = 1$ for illustrative purposes.

network) or s/he may assume that the union of the networks is the observable proxy to be used. It is easy to verify that, in both cases, the estimate would be biased except if, by chance, the true network configuration is correctly guessed. Instead, the model of Section 2 can be used to estimate which network is the relevant one. The parameters would be retrieved from the observation of E, by solving:

$$\left(\hat{\rho}_1, \hat{\rho}_2, \hat{\lambda}\right) \in argmin_{\rho,\kappa,\lambda} \sum_{l=1}^{m} \left[\begin{array}{c} \left(E_{l,G} - a_G - \rho_1^{1+\lambda}(E_{l,B})^{1+\lambda} - \rho_2^{1+\lambda}(E_{l,G})^{1+\lambda}\right)^2 \\ + \left(E_{l,B} - a_B - \rho_1^{1+\lambda}(E_{l,G})^{1+\lambda} - \rho_2^{1+\lambda}(E_{l,B})^{1+\lambda}\right)^2 \end{array} \right] \quad (16)$$

where $E_{-l.t}$ is the effectiveness of a member of type t who is not in group l. Once again we have a three-dimensional problem that can be perfectly identified using the observables $E^* = (E_1^*, ..., E_4^*)$.

4 Estimation

4.1 Model specification

Assume we observe data from $\bar{r}$ Congresses ($r = \{1, ..., \overline{r}\}$), each comprised of n Congress members and associated with an endogenous and unobserved network $G_r = (g_{i,j,r})$. Each legislator i in Congress r is characterized by a level of legislative effectiveness $E_{i,r}$ and a vector of characteristics. We assume there are L observable characteristics and denote them as $X_{i,r} = (X_{i,1,r}, ..., X_{i,l,r}, ..., X_{i,L,r})$. We also assume that there is an observed adjacency matrix linking legislators that may be relevant in the formation of the true network G and denote it with $H_r = (h_{i,j,r})_{i,j \in \mathcal{N}}$.[37] In the following, H_r will be the alumni network, in which $h_{i,j,r} = 1$ if i and j

[37]Multiple observed networks can also be included as input in the model.

have attended the same educational institution in overlapping periods, and $h_{i,j,r} = 0$ otherwise.[38]

Propositions 1 shows that, in equilibrium, effectiveness solves (12). To bring this system of equations to the data, we assume that $\varepsilon_{i,r} = X_{i,r}\beta + \zeta_r + \epsilon_{i,r}$, where β is a $L \times 1$ dimensional vector of parameters, ζ_r is a Congress fixed effect and $\epsilon_{i,r}$ is a random variable with zero mean. We therefore have:

$$E_{i,r} = \varphi \sum\nolimits_{l \in r} \left(\theta_{i,l,r} E_{l,r}\right)^{1+\lambda} \; + X_{i,r}\beta + \zeta_r + \epsilon_{i,r}, \tag{17}$$

where $\varphi = \alpha^{\lambda} \left(\delta\right)^{1+\lambda}$. In (17), the terms $\theta_{i,j,r}$, measuring how costly it is for i to form a link with j, are modeled as random realizations from a logistic function:

$$P\left(\theta_{i,j,r} \left| \chi_{i,j,r}\right.\right) = \left(\frac{e^{\chi_{i,j,r}}}{1 + e^{\chi_{i,j,r}}}\right)^{\theta_{i,j,r}} \left(\frac{1}{1 + e^{\chi_{i,j,r}}}\right)^{1-\theta_{i,j,r}}, \tag{18}$$

with:

$$\chi_{i,j,r} = \iota + \gamma h_{i,j,r} + \sum\nolimits_{l} g(X_{i,l,r}, X_{j,l,r})\psi_l. \tag{19}$$

In (19) ι is a constant, γ and ψ_l are parameters to be estimated, and $g(\cdot,\cdot)$ is a distance function. The value assumed by $\chi_{i,j,r}$ depends on the position of i and j on a known adjacency matrix (for example whether they attended the same school in the same period), and on the distance between i and j in terms of observable characteristics as measured by $g(\cdot,\cdot)$.[39] The specification in (19) therefore allows the cost of forming a link to be random but also to depend on the affinity of the legislators, thus capturing the possibility of homophily.[40]

Party affiliation enters in two ways in the specification of the model presented above. First, the model allows for party affiliation to affect the cost of forming a link. Formally this is done by having a variable $X_{i,j,r}$ in (18) that is equal to one if i and j belong to the same party in Congress r. Party affiliation can, however, also directly affect the benefit of forming a link j by directly entering j's effectiveness. This is captured by including a dummy variable $X_{i,r}$ in (17) identifying party affiliation. We discuss in greater detail our empirical specification in Section 5.

[38] This network has been shown to be relevant as a proxy of social connectedness in Congress by Cohen and Malloy [2014], Battaglini and Patacchini [2018], and Battaglini et al. [2020b]. While we will not assume it to be necessarily relevant, we will use it as input of the analysis as described below. We will discuss this network in greater detail in Section 5.1.

[39] If one observes some links and intends to take these observed links as exogenous, this can be integrated in the model in two ways: (i) by simply replacing the appropriate observed $g_{i,j,r}$ in equation (10) instead of the values in (11); or (ii) by setting the parameter of these exogenous connections in $\chi_{i,j,r}$ in equation (19) equal to infinity. Using this latter approach, one can include multiple exogenous networks and still have an endogenous component.

[40] In our empirical application, we set $g(\cdot,\cdot)$ to be equal to one if both nodes have the same value when the variable is binary, or belong to the same quartile of the distribution of the variable when the variable is not binary. Other functional forms, like the absolute value of differences between individual characteristics, can be used. The logistic function is one of the most popular in dyadic link formation (see for example Graham [2017]). Alternative functional forms and specifications can be used. The estimation method proposed in the next section does not impose particular limitations on these choices.

4.2 Approximate Bayesian Computation (ABC)

To understand our estimation approach, it is first useful to start from the standard estimation approach in Bayesian econometrics. The Metropolis-Hasting algorithm (see Metropolis et al. [1953], Hastings [1970]) is as follows:

A1. Starting from an initial vector of parameters ω, propose a move to ω' according to a transition kernel $q(\omega \to \omega')$.

A2. Calculate $h = \min\left(1, \frac{p(\mathbf{E}|\omega')\pi(\omega')q(\omega' \to \omega)}{p(\mathbf{E}|\omega)\pi(\omega)q(\omega \to \omega')}\right)$, where $p(\mathbf{E}|\omega')$ is the probability of observing $\mathbf{E}$ given ω' in the model and $\pi(.)$ is the prior.

A3. Move to ω' with probability h, else remain at ω; go to the first step.

Under suitable regularity conditions, the limiting stationary distribution of the chain described above is equal to the conditional posterior distribution $p(\omega|\mathbf{E})$. This approach, however, is impossible in our model because an explicit formula for the likelihood function $p(\mathbf{E}|\omega)$ is not available. This problem is not uncommon in complex environments such as ours and it is indeed typical in genetics and evolutionary biology (see Weiss and Haeseler [1998] for instance). Approximate Bayesian computation (ABC) methods allow us to bypass the evaluation of the likelihood function via simulations. Marjoram et al. [2003] has proposed the following algorithm to recover $p(\omega|\mathbf{E})$:

B1. Starting from an initial vector of parameters ω, propose a move to ω' according to a transition kernel $q(\omega \to \omega')$.

B2. Generate $\mathbf{E}'$ using the model with parameters ω'.

B3. If $\varrho(\mathbf{E}', \mathbf{E}) < \nu$, proceed to the next step otherwise return to the first step. Here, $\mathbf{E}$ is the observed vector, $\varrho(\mathbf{E}', \mathbf{E})$ is a norm between $\mathbf{E}'$ and $\mathbf{E}$, and ν is a tolerance parameter.

B4. Calculate $h = \min\left(1, \frac{\pi(\omega')q(\omega' \to \omega)}{\pi(\omega)q(\omega \to \omega')}\right)$.

B5. Move to ω' with probability h, else remain at ω; go to the first step.

This algorithm generates a Markov chain that has a limiting stationary distribution equal to the posterior $\Pr(\omega|\varrho(E', E) < \nu)$. The true conditional distribution $p(\omega|\mathbf{E})$, therefore, coincides with the limit $\lim_{\nu \to 0} \Pr(\omega|\varrho(E', E) < \nu)$.

In a high-dimensional problem as ours, the algorithm above may be slow to execute, since the evaluation in step B2 is computationally intensive: generating $\mathbf{E}$ requires solving for a system of nonlinear equations with hundreds of variables. With the help of our theoretical model, however, we can bypass this step. The idea is to exploit the sharp theoretical characterization of Proposition 1 to directly evaluate how well the proposed vector of parameters ω' solves the equilibrium condition (17) *given* the observed vector $\mathbf{E}$. From (17) define the vector:

$$z(\mathbf{E},\omega) = \mathbf{E} - \varepsilon - \varphi \cdot \Theta \cdot \mathbf{D}(\mathbf{E},\lambda),$$

where $\mathbf{E} = (E_1, ..., E_n)'$ is the observed vector of empirical efficacies,

$$\mathbf{D}(\mathbf{x},\lambda) = ((x_1)^{1+\lambda}, ..., (x_n)^{1+\lambda})',$$

and Θ is a $n \times n$ matrix with i,j element equal to $(\theta_{i,j})^{1+\lambda}$. Let $\varrho(z(\mathbf{E},\omega))$ be the norm of this vector. Our modified algorithm is:

C1. Starting from an initial vector of parameters ω, propose a move to ω' according to a transition kernel $q(\omega \to \omega')$.

C2. If $\varrho(z(\mathbf{E},\omega')) < \nu$, proceed to the next step; otherwise return to the first step.

C3. Calculate $h = \min\left(1, \frac{\pi(\omega')q(\omega' \to \omega)}{\pi(\omega)q(\omega \to \omega')}\right)$.

C4. Move to ω' with probability h, else remain at ω; go to the first step.

The logic behind the modified algorithm C is as follows. Because the equilibrium is characterized by $z(E,\omega) = 0$, we have $E = z^{-1}(0,\omega)$. As a result, requiring $E = \widetilde{E}$ is the same as $z^{-1}(0,\omega) = \widetilde{E}$, or $z(\widetilde{E},\omega) = 0$. This implies that, for each vector of parameters along the Monte Carlo Markov chain (MCMC), we now only need to check that the system $z(\widetilde{E},\omega) = 0$ is close to a zero: $\left\|z(\widetilde{E},\omega)\right\| \leq \varepsilon$ for the ε we use as tolerance. This implies we no longer need to solve nonlinear equations. The properties of this algorithm are characterized in the following proposition.

Proposition 3. *The stationary distribution of the Markov process described by Algorithm C is* $\Pr(\omega | \mathcal{D}_\nu)$, *where* $\mathcal{D}_\nu = \{\omega | \varrho(z(\mathbf{E},\omega)) \leq \nu\}$.

It follows from Proposition 3 that, under the assumption that the model is well specified, $\Pr(\omega | \mathbf{E}) = \lim_{\nu \to 0} \Pr(\omega | \mathcal{D}_\nu)$. That is, the true conditional distribution of the parameters, given the evidence E, coincides with the limit of the stationary distribution as we reduce the tolerance parameter ν to zero. The details of how our ABC algorithm is implemented in practice are described in Section A.3 of the online appendix.

With respect to the previous work in economics that has used the ABC approach, the method presented above is distinctive for two reasons, both of which are derived from the sharp analytical characterization of the equilibrium in (17). First, the condition $\varrho(z(\mathbf{E},\omega))=0$ fully characterizes the equilibrium: this allows us to avoid relying on a partial and arbitrary set of "summary statistics" obtained by simulation to estimate the likelihood function and thus the posterior on the parameters, a problem that is very common in the literature and generally unsolved.[41] Second, our characterization allows us to avoid solving the model in the MCMC, relying instead on the direct evaluation of $\|\varrho(z(\mathbf{E},\omega))\|$ given E.

The analysis presented above and Proposition 3 highlight what distinguishes our approach, which is based on a structural model, from a more data-driven approach that avoids modeling the network formation stage. The observable network G that we aim to estimate is a high dimensional object (an $n \times n$ matrix) compared to the observable outcome (n dimensional vector). Using exclusively this type of data, we would need to observe the effect of the network on $\mathbf{E}$ for many periods. This would not be realistic in many applications (including the case of the U.S. Congress) because networks rarely remain constant for long periods of time; and it is highly unlikely that output (i.e. effectiveness in our application) can be accurately measured using high frequency data (such as at a weekly or monthly basis). Our approach, instead, uses the model to reduce the dimensionality of the problem, which allows us to focus only on the estimation of the model parameters as opposed to focusing on G directly. Proposition 3 clarifies the output of this procedure: we obtain the posterior distribution of the parameters, given the observable data and the model.[42]

We conclude this section with a comment on identification of the model in our Bayesian analysis. Provided that it is proper, the posterior $\Pr(\omega\,|\mathcal{D}_\varepsilon)$ is always well defined and it incorporates all information in $\mathbf{E}$ given the model (Lindley [1971], Kaas et al. [1998]). In addition, the interpretation of this posterior is straightforward. Bayes rule allows us to use the observations to update the probabilities of the events associated to the sigma algebra of the minimal sufficient parameter space.[43] If the parameter space in the Bayesian model is not minimal, the conditional probabilities on events associated to their finer sigma algebra are naturally updated relying on their prior probabilities. A parameter is therefore not identified in Bayesian theory only if its prior distribution is not revised through the information brought by the data, so that the conditional posterior and conditional prior distribution are the same (Florens and Simoni [2011]). In the next section, we show with simulations that the Bayesian procedure described above provides accurate estimates of all parameters for the type of datasets that we use in the empirical application in Section 5. Regarding the empirical analysis of Section 5, Figures A.9 - A.11 in the supplementary

[41]See Konig [2016] and Boucher [2020] for alternative criteria on how to select these statistics in the absence of a complete characterization.

[42]We should note that in the type of economic environment we are studying (in which the social network is unobservable) it is unavoidable to rely on a model because we need to establish a theoretical link between E and G to use the former in the estimation of the latter. The approaches described in the discussion of the related literature based on LASSO techniques typically assume a linear model linking E to G. What distinguishes our analysis is that we are providing a complete microfoundation of the link between E and G. This, together with Proposition 3, helps make the interpretation of the posterior generated by our approach more transparent.

[43]A Bayesian model can be seen as a statistical model on which the parameter space A is endowed with a probability measure on $A, \mathcal{A}$, where $\mathcal{A}$ is the sigma algebra of A. The parameter space $\mathcal{A}$ is sufficient if it is sufficient to describe the sampling process. The minimal sufficient parameter $\mathcal{A}^*$ is the intersection of all the sufficient parameter $\mathcal{A}$. See Florens et al. [1990] for a reference.

online appendix show that there is significant updating of information in the chains, which results in posteriors that are markedly different from the priors.

4.3 Monte Carlo simulations

In this section, we use simulations to investigate the performance of our estimation approach. We first propose two examples to illustrate its ability to recover the underlying structural parameters and key features of the network topology. We then systematically study how the performance of our method changes as we vary important features of the setting (i.e. the number of periods for which we observe the outcomes for the same network, the size and sparsity of the networks, the elasticity of network formation, and others), and of the algorithm used in the implementation of the ABC.

4.3.1 Two benchmarks

The alumni network. There are a number of parameters determining the shape of the social network in our model. An important ingredient is the observed adjacency matrix that enters in the cost function (3), or in other words, the observed factors affecting the cost of establishing a link from legislator i to legislator j. In our first set of simulations, we use an observed network as a basis to simulate $H = (h_{i,j})$. Specifically, we generate H using the alumni network of the politicians of the U.S. Congress, the same network that we will use in our empirical application. For the 111th Congress, we randomly select $n = 200$ legislators and set the observed adjacency networks $h_{i,j} = 1$ if i and j graduated from the same school and zero otherwise.[44]

We use a given set of parameters $(\iota, \gamma, \psi_l, \varphi, \lambda, \beta)$, randomly generated characteristics X, and H as inputs in (17)-(19) to derive the observed vector of effectiveness $\mathbf{E}$ and the true network $G = (g_{i,j})$. We simulate an environment in which the vectors of effectiveness in five Congresses are observed, which corresponds to simulating five vectors $\mathbf{E}$. The parameters are calibrated so that the resulting network matches the degree and clustering coefficients in the cosponsorship network.[45] The estimation procedure consists of using the simulated $\mathbf{E}$, H, and $\mathbf{X}$ to estimate $(\iota, \gamma, \psi_l, \varphi, \lambda, \beta)$, and thus the unobserved G. Because we know the true $(\iota, \gamma, \psi_l, \varphi, \lambda, \beta)$ and G, we can evaluate the performance of the estimation technique.

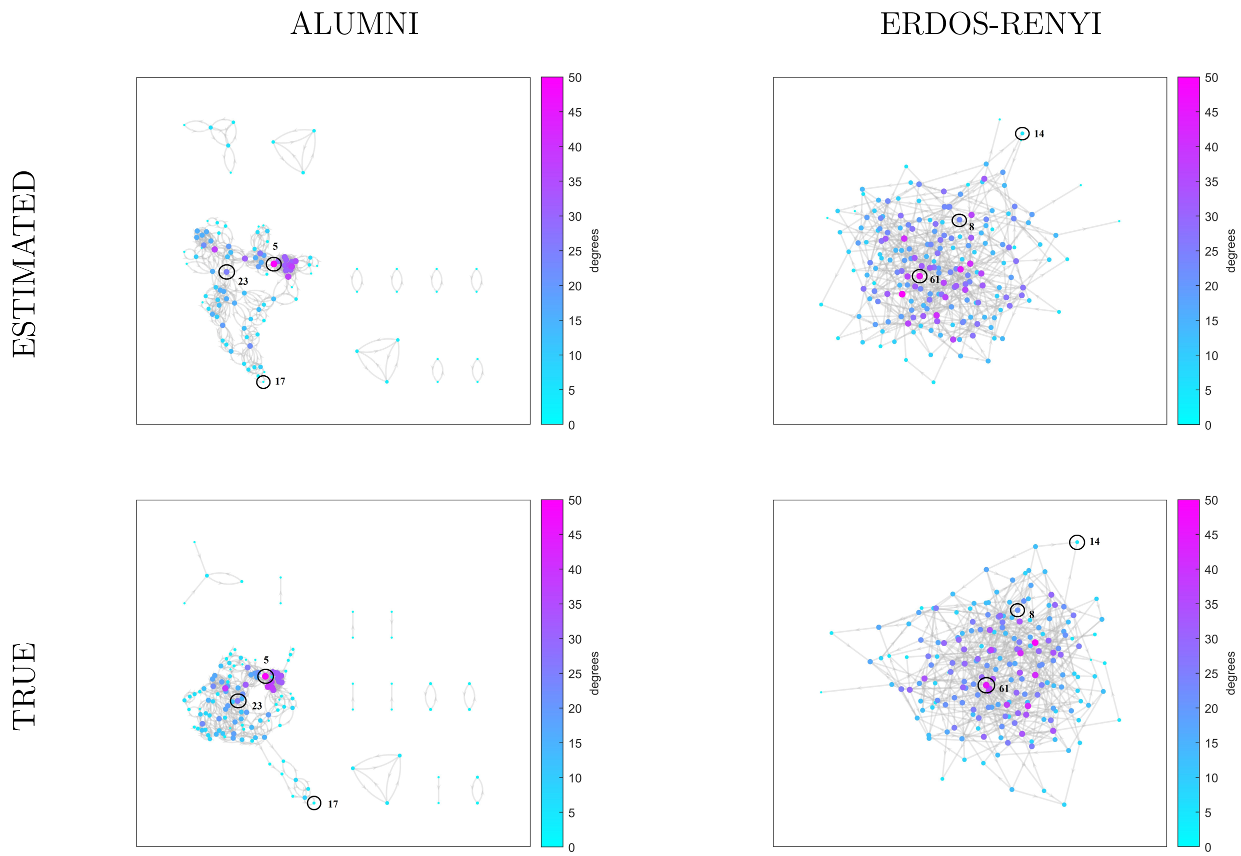
To illustrate the performance of the network estimation, in all of our simulation experiments we randomly select one of the five simulated networks that is generated at the 10,000th simulation from the ergotic distribution of the first of the MCMCs used for the estimation after a burn-in period of 10,000 iterations.[46] In Figure 3, we show a graphical comparison of the true network (the

[44]The 111th Congress is chosen at random as an example. Similar findings are obtained with other Congresses.

[45]Further details regarding the simulation setup are presented in Section 8.5 in the appendix.

[46]Given that the posteriors of estimated parameters are highly concentrated around the true values, this network is very close to one constructed using median values of the parameters from the posterior distribution, but faster to compute. See Section A.3 for the details of the implementation. Similar evidence is obtained when using any

Figure 3: NETWORK ESTIMATION
- QUALITATIVE EVIDENCE -



NOTES. Upper panels represent the estimated network, and lower panels represent the true network. Estimated networks are generated as described in Section 8.5. The cost of forming a link depends on alumni connections in the left panel and on the Erdos-Renyi network in the right panel. The algorithm uses attractive forces between adjacent nodes and repulsive forces between distant nodes in the network. See Fruchterman and Reingold [1991] for more details. The size of the nodes is proportional to their degree. Three random nodes are drawn and highlighted with circles.

lower panel) versus the estimated network (the upper panel). In the left panel, the true network is the alumni network.[47]

The pictures show that the estimated and true networks are, naturally, not identical but remarkably similar in terms of topological structure. In our model, even taking the parameters as given, a network depends on deterministic factors and random components affecting the cost of forming a link $\theta_{i,j}$. Dyads, triads and dense clusters are well represented in the estimated network, and nodes appear in their true topological position. To highlight the topological similarities, we mark three specific nodes, one in the center, one in the periphery, and one in the extreme periphery of a component of the network. Their respective positions in the true network are preserved in the estimated network. More formal evidence of the goodness of fit of our estimation method is provided by the ROC curve for the true positive rate of estimated links in Figure A.2 in the online

other network of the simulated five, or other draws, or other chains.

[47]The right panel will be described in the next section.

appendix.[48] Figure A.2 shows that almost all of the links are correctly predicted.

As further evidence of the good performance of the method, in Figure 4, we plot the values of individual network statistics for each node (betweenness, eigenvalue, closeness, clustering, and in- and out-degree).[49] Each point corresponds to a node of the network, with the estimated network statistic on the X-axis, the true network statistic on the Y-axis, and the bisector drawn in red. The plots reveal that the data points lie close to the red line, again showing the ability of the ABC procedure to precisely estimate the position of each node in the network. Finally, panel (a) of Table A.1 in the online appendix shows aggregate network statistics in the simulated and true networks (density, assortativity, closeness, betweenness, degree, and clustering). Perhaps unsurprising at this stage, the table confirms that the estimated values are quite similar to the true values in the model.

The Erdos-Renyi network. Our second benchmark is a simulation in which $h_{i,j}$ is generated using the Erdos-Renyi model in which the probability of each link is equal to p, a standard benchmark in the theory of networks. While this setup is less realistic than the first one, it will prove useful in the next subsection for the comparative statics exercise. As in the previous section, here we also first generate $r = 5$ observations with $n = 200$ legislators in each; we then use the realized effectiveness alone to estimate the parameters and the unobserved network G. A similar analysis as in the previous section is presented in the right panels of Figure 3, Figure A.3, and panel (b) of Table A.1. Results are qualitatively very similar as in the simulations based on the alumni networks described above.

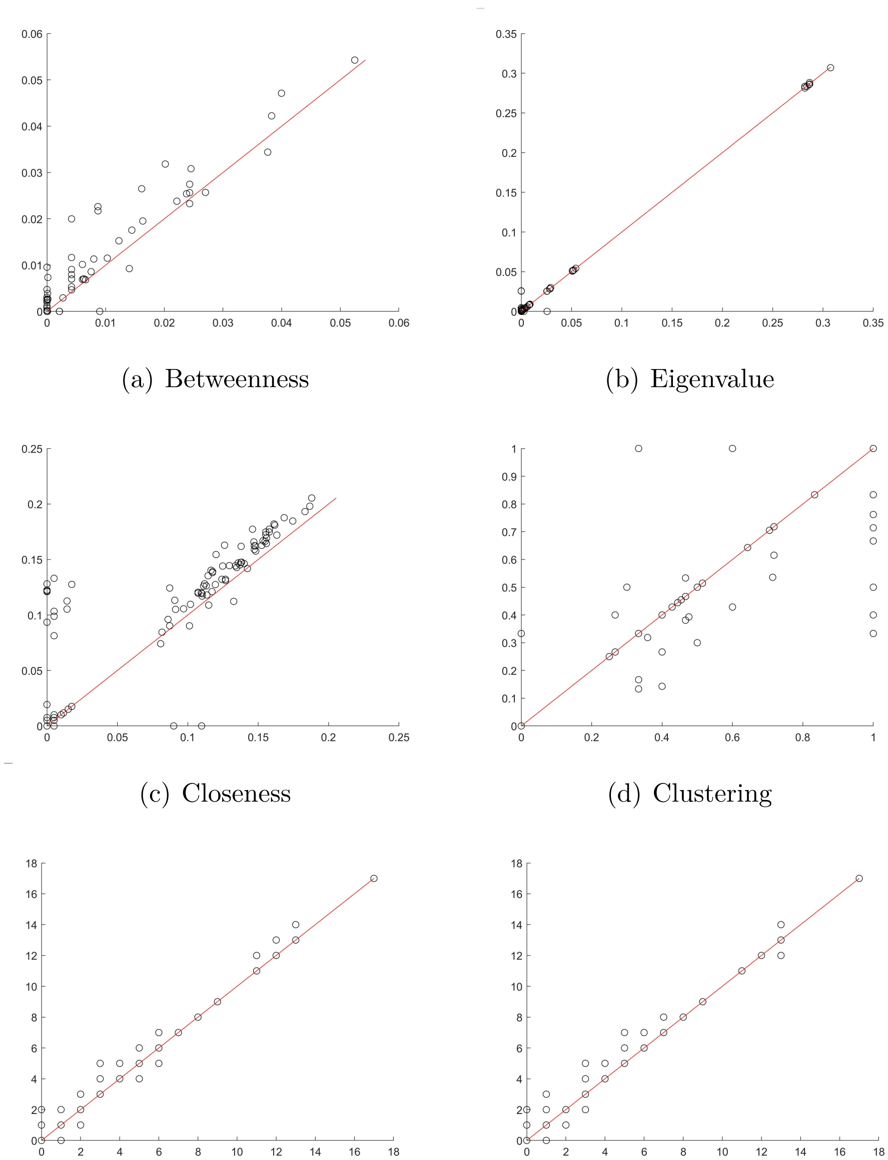
4.3.2 Sensitivity Analysis

We now explore the performance of our method as we change key features of the economic environment and of the algorithm used to implement the ABC procedure. In terms of the environment, we change the number of observations r, the number of legislators n, the sparsity of the network as measured by $p = \#links/[(n-1)n]$, the elasticity of link formation λ, and other important aspects of the network's topological structure. In terms of the algorithm, we change the number of MCMCs used to construct the posterior distribution of the parameters. We not only show how our method performs in terms of parameter estimation, but also how it performs in terms of the estimation of node-level characteristics. The comparative static exercises presented below are done using the Erdos-Renyi network as the true network. For completeness, the same exercises and figures reported here are also reported for the alumni network in the online appendix Figures A.5, A.6 and A.7. The results are qualitatively very similar for the two cases.

[48] The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR) at various thresholds. It can also be interpreted as a plot of the power as a function of the Type I error of the decision rule. The closer the curve is to the upper right contour of the box, the higher the TPR, the lower the FPR, and thus the better is the prediction of links.

[49] See Newman [2010] for a formal definition of these centrality measures.

Figure 4: NODE-LEVEL STATISTICS
- ESTIMATED VS TRUE NETWORK -
- *ALUMNI 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 alumni connections. 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.

Number of observations. We start by varying the number of observations r from 1 to 10 while keeping the number of nodes in the network constant at 200. In the application that we will study in Section 5, this corresponds to changing the number of Congresses that are used in the estimation. The left panels of Figures 5 and 6 report the distribution of the estimation bias along the Markov chain. Each box plot represents the difference between the real and the estimated parameter (Figure 5) and the node-level characteristics (Figure 6) when r varies.[50] The plots show a clear reduction in the mean and variance of the distribution as we increase r. The important point here is that just two observations (i.e. $r = 2$) are sufficient to obtain estimates with a negligible bias; and a single observation (i.e. $r = 1$) is sufficient for obtaining a good match of the key centrality measures. This is an important difference with respect to the literature attempting to estimate sparse networks using repeated observations with LASSO techniques. While these approaches do not require a formal model of endogenous network formation and attempt to directly estimate the network links (rather than the parameters of a model of network formation), they require repeated observations from the same network (easily over 20 times) to obtain reliable estimates.

Network size. We have simulated networks for $n = 50, 100, 150,$ and 200 nodes. The plots on the center of Figure 5 show that, for each parameter, the distance from the true value and its dispersion converges to zero as n increases. It is interesting to observe that $n = 150$ is already sufficient to have distances highly concentrated around zero. The plots on the center of Figure 6 present box plots where each box represents the difference between the real and estimated node-level statistics. Here as well, the graphs show that the distributions concentrate around zero as n increases. This simulation exercise provides ample evidence that our methodology is able to estimate social interactions in environments similar to the environment of our empirical investigation, where we will deal with five Congresses of about 400 politicians.

Number of MCMCs. Our estimation procedure allows for the use of multiple MCMCs. The use of more chains enables us to shorten their length and explore the parameter space more comprehensively in a shorter amount of time. We have estimated our model with a number of chains $c = 1, 4, 8$ and 16. The plots on the right of Figure 5 show that the distance from the true value and its dispersion converges to zero as c increases for each parameter. While we obtain estimates close to the true value even with a small number of chains, the performance of the algorithm improves in all dimensions as we increase them. This is confirmed by Figure 6, which presents a box plot where each box represents the difference between the real and the estimated node-level statistics. Here too, the graphs on the left show that the distributions concentrate around zero. The performance is already remarkably good starting at eight chains. Based on the

[50] The box plots for the parameters are constructed using the ergodic distribution generated by 10,000 draws from our Markov chains after a burn-in period of 10,000 iterations each. The distributions of node-level characteristics are constructed using the network generated as described above. In Figure 6 we report closeness, betweenness and eigenvector centralities, in Figure A.4 in the online appendix we also report clustering, in- and out-degree. We normalize the realizations by subtracting the true values from each of them, thus centering the distributions at zero in all cases.

evidences from these experiments, we use 16 chains in the rest of simulation exercises and in our empirical application below.

Network topology. We now show the performance of our methodology under alternative network topologies. We consider three cases: the alumni network, the Erdos-Renyi network described above, and a circular network.[51] Figure A.8 in the online appendix displays the true adjacency matrix, the associated estimated matrix using our methodology, and their difference when the true network is simulated under different rules. Remarkably, the estimated linking structure closely follows the changes of the true network in all of the different network structures.

Other exercises. To conclude, we extensively explore the performance of our estimator along two other important dimensions: the density of the network and the elasticity of network formation as measured by λ.[52] For both measures, we consider two different network topologies: the topology of the alumni connections and that of the Erdos-Renyi network. The analysis is presented in Section A.4 of the online appendix. With respect to network density, we find 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. With respect to λ, the results reveal no systematic pattern across its values, and that the distributions are mainly concentrated around zero for all values of λ and with similar dispersion. We conclude that the performance of our methodology does not hinge on a particular value of λ.

5 Evidence from the U.S. Congress

5.1 Data description

We measure each Congress member's legislative performance using the Legislative Effectiveness Scores (LESs) for members of the U.S. House of Representatives, which were developed by Volden and Wiseman [2014]. Each member's score is based on how many bills each legislator introduces, as well as how many of those bills receive action in the committees, are approved at the committee level, receive action of the floor of the House, pass the House, and ultimately become law. In addition, we also analyze effectiveness separately at each stage of the legislative process. Data are available online from the Legislative Effectiveness Project (http://www.thelawmakers.org).[53] We use information from five recent election cycles: the 109th Congress (election cycle 2004) to the

[51] In a "circular network," i is linked to agents $i+j$ for $j \leq z_i$, where z_i is an independent realization from a uniform distribution with mean m.

[52] 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.

[53] Volden et al. [2013] have used this data to explore the legislative effectiveness of women for the 93rd - 110th Congresses. A similar score, Health ILESs, was proposed by Volden and Wiseman [2011] to examine which House members have been most successful at advancing health care bills for the 93rd to 110th Congresses. The scores are not normalized by the total number of bills proposed by all politicians to avoid spurious correlations.

113th Congress (election cycle 2012). Consistent with existing theories of congressional politics, Volden and Wiseman [2014] argue that legislative effectiveness is a function of innate abilities, a cultivated set of skills, and institutional positions.

Our analysis considers all of the legislator characteristics indicated in the Legislative Effectiveness Project. Specifically, the control set $X_{i,r}$ in model (17) includes the number of years spent in Congress and its squared term, margin of victory and its squared term, DW-ideology, the state in which the Congress member was elected, the size of the state congressional delegation, party, chairmanship, majority and minority party leadership, whether the representative is a member of the most powerful committees, previous legislative experience, gender, and race. In addition to these factors, we include the age of the legislators as a demographic characteristic, and their main area of policy interest. The inclusion of the policy area of interest as control variable is especially important in our context since an increase in salience of a topic in a legislature may result in greater effectiveness of the legislators promoting bills in that policy area. For each Congress member, we identify the main policy interest in the following way. We retrieve the data provided by the Congressional Bills Project (http://congressionalbills.org), which categorizes the bills sponsored and cosponsored by each Congress member using the policy topic coding system provided by the Policy Agendas Project (PAP) (www.comparativeagendas.net/us). For each Congress member i, we count the bills where the Congress member i was an original sponsor or an original cosponsor in each policy subtopic, and identify her/his most recurrent policy subtopic.[54]A precise definition of the policy content related to each PAP category is available at https://www.comparativeagendas.net/pages/master-codebook.

The main legislator characteristics that may affect the cost of forming a link (θ_{ij}) are also included in the network formation model (18)-(19). This set includes similarities in terms of party, leadership, gender, ethnicity, seniority, age, area of policy interest, and state. Because the cost of forming links may be closer to zero for more senior members, for which connections are already formed, we also include seniority as an individual-level variable.[55]

It is important to understand that in our model network formation is affected by both the variables included in (18)-(19) and by the variables in (17). The variables in (18)-(19) affect the cost of forming the link because they contribute to the determination of the $\theta_{i,j}$s that define (3). The variables in (17), instead, affect the marginal benefit of establishing a link because they

[54]Specifically, we order the interest of a legislator for different subtopics by first ranking the subtopic category on which he/she sponsored or cosponsored the highest number of bills. The subtopic category corresponding to the highest ranking is considered the most relevant to the legislator's agenda. When multiple subtopic categories are ranked first, because the legislator sponsored or cosponsored the same number of bills for these categories, we consider each of these categories to be relevant for the legislator's agenda. We then construct a dummy variable for each subtopic with value equal to one if the politician (co)sponsored the majority of bills in that subtopic or zero otherwise. In the Congresses under analysis, there are 54 subtopics that are the main policy area of interest for at least one politician.

[55]We do not include the full set of controls in order to decrease the number of parameters to be estimated, thus easing the computational burden of the Bayesian estimation techniques.

contribute to the determination of the equilibrium effectiveness of the lawmakers defined in (17).

To construct the alumni network, we extract information on the educational institutions attended by the Congress members using the Biographical Directory of the United States Congress, which is available online (http://bioguide.Congress.gov/biosearch/biosearch.asp).[56] We assume that a tie exists between two Congress members if they graduated from the same institution within four years of each other. Because many legislators hold a primary and a secondary degree (typically a JD or an MBA), this construction gives us a rich network of direct and indirect links. Additional details on the construction of the alumni network can be found in Battaglini and Patacchini [2018]. The Biographical Directory of the United States Congress is also used to retrieve the information about the date of birth of the Congress members which is used to calculate their age. Table A.2 in the online appendix provides a detailed description of the variables used in this study, together with summary statistics for our sample.

5.2 Empirical findings

Using the procedure described in the previous sections, we now proceed with the estimation of model (17)-(19) using data from the 109th Congress (election cycle 2004) to the 113th Congress (election cycle 2012). Table 1 and Table 2 present the median value of the posterior distributions of the estimates of the key parameters. Table 1 shows the parameters φ and λ, and the β that enter in the definition of ε_i in equation (1). Table 2 shows the median values of the estimates of the parameters shaping the network formation: ψ, γ, and ι. These tables also report the probability of observing a value greater than zero (p-value at zero) on each parameter's estimated posterior distribution. A p-value equal to one indicates that the support of the entire distribution is strictly greater than zero, whereas a p-value equal to zero means that the support of the entire distribution is less than or equal to zero. The posterior distributions of all of the key parameters of the model are presented in Figures A.9-A.11 in the online appendix.

The size of the social spillovers. We start by discussing the key parameters concerning the size of the social spillover and their impact on effectiveness: φ and λ. These parameters are found to be positive and statistically significant. The positive and statistically significant value of φ measures the social spillovers in the model. A one percent increase in the social connectedness of i, as measured by s_i in (1), induces a 0.80 percent increase in effectiveness E_i.[57] For λ, recall that $\lambda = \epsilon_{g_{i,j},E_j}$, where $\epsilon_{g_{i,j},E_j}$ is the elasticity of $g_{i,j}$ with respect to E_j. When $\lambda = 0$ the analysis reduces to a model in which the connections $g_{i,j}$ are assumed to be known, exogenous and equal

[56] We use high schools and academic institutions attended for both undergraduate and graduate degrees. In dealing with multiple campuses, we match each satellite campus as a separate university (e.g., University of California at Los Angeles, San Diego, and Berkeley are treated as separate universities). We match specialized schools (e.g., law schools) to the larger university.

[57] This value is the median of the estimated posterior distribution of α (the elasticity of effectiveness of social connectedness in equation (1)), which is reported in Figure A.10 in the online appendix.

to $\theta_{i,j}$. The fact that λ is significantly greater than zero highlights the importance of the fact that the network is endogenous: an increase in the effectiveness of j by 1 percent leads to a marginal increase in $g_{i,j}$ of 0.60 percent, ceteris paribus.

In Table 3, we investigate whether network effects are more important in the early or late stages of the legislative process. Following the classification presented by Volden and Wiseman [2014], we decompose the LES score using the following bill categories: bills proposed in the House of Representatives (BILL), bills that received any action in committee and bills that received some action beyond the committee stage but have not passed in the House (AIC-ABC), and bills that have passed in the House (including those that have become law) (PASS-LAW). Column (1) reports our baseline estimation. In Column (2), (3), and (4), the dependent variables are respectively BILL, AIC-ABC and PASS-LAW. The results in Table 3 show that network effects are positive and significant for all stages of the legislative process, especially at the beginning.

Fit of the model. The findings discussed above should be contrasted with two benchmarks. The first is the standard model that ignores network effects. This model is typical of the traditional literature on legislative effectiveness that has focused on the individual characteristics determining effectiveness while ignoring social spillovers.[58] We refer to this as the *no-network model*. The second benchmark is the case in which legislative efficacy depends on social connections, but the elasticity of network formation λ is zero, so the connections are exogenous and equal to $\theta_{i,j}$. We refer to this as the *exogenous network model*. Both of these models can be seen as nested special cases of our more general model of endogenous network formation. The no-network model corresponds to our model when we impose $\varphi = 0$, so that social networks are assumed to be nonexistent. Comparing this model with ours, therefore, allows us to assess whether it is important to consider social connections. The exogenous network model corresponds to our model in which we allow for a positive φ, but we impose $\lambda = 0$. Comparing this model with ours not only allows us to understand the importance of social connections, but also whether it is important to model endogeneity.

Table 4 collects the estimates. The first two columns report the results from the no-network model and the exogenous network model models, respectively. The last column reproduces for convenience the estimates from Table 1, when endogenous network effects are allowed. We want to formally check whether the model fit improves across columns. To accomplish this, we use different goodness-of-fit measures. Because both the no-network and the exogenous network models are sequentially nested in our model, a standard partial F-test is a natural first approach.[59] Results

[58] Empirical analysis of this type have been presented by, among others, Jeydel and Taylor [2003], Padro I Miquel and Snyder [2006], Volden and Wiseman [2014].

[59] Let RRS_1 define the residual sum of squares of the unrestricted model and p_1 the number of parameters. Let RRS_2 the residual sum of squares of the restricted model, and p_2 the number of parameters. The partial F-test statistic $F = [(RRS_1 - RRS_2)/p_1 - p_2]/(RRS_1)/n - p_1$ will have an F distribution with $(p_1 - p_2, n - p_1)$ degrees of freedom if the residuals are normally distributed. The Kolmogorov Smirnov test cannot reject the hypothesis

are presented in Columns (2) and (3) of Table 4. The F-test rejects the hypothesis that the model with exogenous network effects does not provide a significantly better fit than the model without network effects, and that the model with endogenous network effects does not improve the fit of the model with exogenous network effects.

Next, we compare the models using three alternative measures of fit based on prediction error: the R^2, the Mean Squared Error (MSE), and the Mean Absolute Simulated Distance (MASD), which is the mean value of the absolute distance between the observed and the simulated data.[60] Table 4 also reveals that the fit improves across columns according to each of these measures.

Finally, we assess the performance of different models using the *Deviance Information Criterion* (DIC). The DIC is a generalization of the Akaike information criterion (AIC), which can be used when the likelihood is not available (see Spiegelhalter et al. [2002]). This measure is used in Bayesian model selection problems where the posterior distributions have been obtained by simulating data with an MCMC, such as in the ABC method. The value of the DIC not only depends on how close the model predictions are to the data, but also on the complexity of the model, with more complex models being penalized. The model that receives the highest support from the data is the one with the lowest value.[61] As it can be seen from Table 4, the DIC reaches

that they are normally distributed (the value of the test is 0.0158, which is below the critical value of 0.0348 at the 1 percent confidence level). The RSS is computed by extracting 500 realizations, evaluated at the median of the posterior distribution of all parameters, and then computing the average value of the predicted vector of outcomes ($\hat{y}$), by generating the random errors and averaging across simulations. In particular, for the endogenous network we set $\theta_{ij} = 1$ if $P\left(\theta_{i,j,r}\left|\chi_{i,j,r}\right.\right) > u$, where u is drawn from a uniform $U(0,1)$. The residuals are then computed subtracting the predicted values of effectiveness from the realized ones.

[60]Because our model is not linear, we use a pseudo-R defined as $R^2_{RG} = 1 - \frac{Q_{fit}-Q_0}{Q_{min}-Q_0}$, where RG stands for "relative gain", Q denotes an objective function to be maximized (or minimized), Q_0 is its worst possible value, Q_{min} is its minimum possible value, and Q_{fit} is the value estimated (Cameron and Trivedi [2005], chapter 8.7). We propose two versions of this indicator. In the first, called pseudo-R^2_1, the objective function is the traditional residual sum of squares RRS. In the second, called pseudo-R^2_2, the objective function is $\varrho\left(z(\mathbf{E},\omega)\right)$, that is the objective function that we use in Algorithm C (see Section 4.2). More precisely, in the first we use a version of the MSE, i.e. $MSE = \sum_{i=1}^{N}(y_i - \hat{y_i})^2/(N-K)$, where $\hat{y_i}$ is the prediction of y_i under the estimated model; and in the second, we set $Q = \sum_{i=1}^{N}|y_i - \hat{y_i}|/(N-K)$. We set Q_{min} equal to zero in both versions, as the minimum value is reached when the simulated data is always equal to real data; and we set Q_0 equal to $\{[max_i(y_i) - min_i(y_i)]/2\}^2$ and $[max_i(y_i) - min_i(y_i)]/2$ in the first and second versions respectively, since our dependent variable is bounded between 0 and 1 and thus we consider the worse prediction equal to half of the data range. Because the values of R^2 increase with the number of parameters K, we use the adjusted versions where the denominator is $(N-K)$. Both measures assume values between 0 and 1. As explained above, to compute $\hat{y}$, we simulate 500 vectors y using the median values of the posterior parameters and randomly extracted random errors, and then compute the average vector across the simulated values. The MASD is calculated using 500 simulated samples.

[61]The DIC is particularly suitable to compare models without likelihoods because posterior predictive distributions and approximations of the deviance can be used to compare different models. Models are favored not only by a good fit, but also by low complexity (similar to AIC). The general form of DIC is $DIC = \bar{D} + p_D$ where $\bar{D}$, the posterior expectation of the deviance, is $E_{\theta|Y}[-2logP(Y|\theta)]$ and p_D (the complexity of a model) is the difference between $\bar{D}$ and the deviance evaluated at a particular point estimate $D(\hat{\theta})$. An example of $\hat{\theta}$ often used in applications is the estimate of the posterior mean of the model parameter. $\bar{D}$ prizes models that produce data close to the observed, while also considering the dispersion of the estimated parameters. p_D penalizes models with large numbers of parameters and high parameters dispersion. We use the procedure in Francois and Laval [2011] to compute the DIC with ABC models. In particular, we consider DIC_1 because it is faster to compute. Specifically, $DIC_1 = \bar{D} + p_D$ where $\bar{D}$ denotes the posterior expectation of the deviance $E_{\theta|Y}[-2logP(Y|\theta)]$ and p_D is the

the lowest values for the model with endogenous network effects with a remarkable gain in fit. The model with endogenous networks is associated with a 17 percent decrease in the DIC with respect to a model with no network effects.

To complete the analysis of the performance of our approach, we compare it to three other models that, while not nested in our model, can be seen as standard alternative ways to control for the endogeneity of the social network. The first is the traditional linear model of social interactions (Spatial Auto Regressive Model, SAR) where an exogenous matrix is used to approximate social connections. A natural candidate in our case is to use the alumni connections. We refer to this as the *SAR Model*. The second is the model suggested by Battaglini et al. [2020b]. This model uses the (observable) cosponsorship network $G^{Cosp} = \left(g_{i,j}^{Cosp}\right)$ as a proxy for social connections but controls for endogeneity with a two step approach *a' la* Heckman. In the first step, each link in the cosponsorship network $g_{i,j}^{Cosp}$ is estimated as a linear function of the alumni network $H = (h_{i,j})$ and the distance between i and j in terms of observable characteristics.[62] This gives us a matrix of residuals $\left(\widehat{\varepsilon}_{i,j}^{Cosp}\right)$ that captures unobserved factors affecting the link between i and j. Battaglini et al. [2020b] characterize conditions under which the residuals can be used as a control for the selection bias and estimate a SAR model with the residuals as an additional regressor. We refer to this as the *Two Step Model I*. The third is a variant on this approach in which in the first step we again estimate the cosponsorship links on the alumni connections and characteristics, and then, in the second step, we use the matrix of predicted links $G^{\hat{Cosp}}$ instead of the matrix of cosponsorship links in the SAR model of effectiveness. We refer to this as the *Two Step Model II*. A limitation of these two approaches is that, in the absence of a model of network formation, there is no obvious way to "estimate" a matrix of observed cosponsorship links with a matrix of alumni connections in the first step. Even if we accept the linearity in the homophily model in the first step, its assumed pairwise nature ignores important structural network characteristics in the estimation of the social network such as the propensity to form ties between individuals sharing one or more contacts, or a different likelihood to form new ties for low- and high-degree nodes. Our model of endogenous network formation incorporates these second order effects because each link in equilibrium reflects the equilibrium levels of effectiveness in the entire network.

Table A.3 compares each of these alternatives with our model.[63] In the last rows, we report

difference between $\bar{D}$ and the deviance evaluated at a particular point estimate $D(\hat{\theta})$. We follow the convention in the literature and set $\hat{\theta}$ equal to the mean of the estimated posterior of the model's parameters. We consider the mean absolute deviance of simulated data from real observation as the main summary statistic, a normal kernel with bandwidth equal to one. The DIC is calculated using 500 simulated samples.

[62] In the cosponsorship network, a link between i and j is established if i cosponsors bills by j. To have a density comparable to the alumni network, we set that a link exists if i cosponsored at least 10 bills proposed by j and zero otherwise, which corresponds to a winsorization of the distribution of cosponsorships at the 99th percentile. We use the functional form of the distance in age, gender and race as in model (19) described in Section 4.1.

[63] We estimate these models with our ABC algorithm. In the SAR Model, we set $\lambda = 0$ and Θ equal to H in equation (17). In the Two Step Model I, we set $\lambda = 0$ and Θ equal to G^{Cosp} in equation (17) and we add the residuals from the first step as a control in the X of (17). In the Two Step Model II, we set $\lambda = 0$ and Θ equal to $G^{\hat{Cosp}}$ in equation (17).

the measures of goodness of fit described above.[64] All of the measures agree that the model with endogenous network formation performs significantly better than the others.[65]

The social network. Column (1) of Table 2 provides insight on the characteristics that matter for social connectedness in the U.S. Congress.[66] As is apparent from Table 2, we find that belonging to the same party has the highest impact on the cost of forming a link. The second most important factor, perhaps unsurprisingly, is the state in which the Congress members are elected. Having an alumni connection is also a significantly positive factor (remarkably, about one-tenth the size of party affiliation). This is in line with previous studies documenting that alumni connections are important predictors of social interactions, including voting behavior and cosponsorships (see Cohen and Malloy [2014], Battaglini and Patacchini [2019], among others). Congress members with longer tenures have a larger number of connections, consistent with the idea that their connection costs are lower because some connections are already formed. We also find that connections are more likely to be established between politicians with similar levels of seniority. Other factors that contribute positively to link formation are similarities in terms of age, state, and policy preferences. Similarities in gender, race, and institutional position (i.e. both serving in a leadership positions) are instead negatively correlated with link formation.

Using the estimates of Table 2, we reconstruct the unobserved social network and study its characteristics. Figure 7 depicts the estimated network for the 111th Congress, using the coefficients of Tables 1 and 2.[67] Three features of the social network uncovered by our analysis are worth highlighting. The first feature is the importance of party affiliation. In Panel (a) we color the nodes denoting Democratic (respectively, Republican) politicians in blue (resp., red). Democrats have a clear higher propensity to link with fellow Democrats, and the same is true for Republicans. This is a feature that we would have underestimated by using other observable adjacency matrices commonly used to study congressional behavior (such as the alumni network or the cosponsorship network) as proxies for the true social network because they do not exhibit such polarization. Figure A.1 compares the estimated network, the alumni network, the committee network, and the cosponsorship network when coloring nodes by party affiliation. The figure indicates that our estimated network is better at capturing polarization than the other networks.

An advantage of our approach is that it not only provides a binary measure of connection, but also an intensive margin, which allows us to identify the strength of interactions. The second

[64] Because these models are not nested, we do not use the F statistics.

[65] Interestingly, the DIC and MASD slightly increase when exogenous network effects are included in the model but decrease significantly when the network is endogenized. This could be explained by the fact that the increase in model complexity is not offset by a significant increase in the fit of the data when an arbitrary exogenous network is considered. Instead, when the network is endogenized and the results from the optimization of the politicians' choices are included, the increase in complexity is more than offset by a better fit.

[66] Column (2) provides similar information in a modified model in which we also allow for unobservable covariates. It will be discussed in detail in Section 6.3.

[67] Qualitatively similar results are obtained with other Congresses, but we omit the analysis of these cases here for brevity.

feature of the social network worth highlighting can be seen in Panel (b) of Figure 7, where the color of the links is proportional to the value of their respective $g_{i,j}$. We see that, besides the high intensity of intraparty connections highlighted above, legislators also develop weaker interparty connections. This is interesting because it confirms a theory that developing "weak ties" with a heterogeneous set of agents (especially, in the case of Congress, with different political ideologies) is essential for advancing bills through the legislative process (see Granovetter [1973] and especially, for the U.S. Congress, Kirkland [2011]).

The final feature we want to highlight is an asymmetry between Democrats and Republicans: the Democrats' cloud is more dispersed in terms of the node's connectivity than the Republican cloud, which features many members with few connections between each other. Democrats have more nodes with a high number of connections, stronger links, and are consistently more effective than Repubblicans. This asymmetry reflects the fact that Democrats were the majority party in the 111th Congress, which explains the higher average level of effectiveness.[68]

We conclude this section with a final note on the identification of the $g_{i,j}$s. One may ask whether the network estimation would be biased if there are unobserved characteristics that affect both link formation and legislative effectiveness. In section 6.3 we extend our model to control for this issue and show that the estimates are robust to this type of concern in our application.

A Comparison between the Estimated and Observed Networks. We complete the discussion of the previous subsection with a more detailed comparison of the estimated network with other known and observable structures of interactions among politicians used in the literature to approximate political interconnectedness. Table 6 reports some network-level statistics computed for the estimated network G and the cosponsorship, committee, and alumni networks. It appears that the density, degree, and closeness in our estimated network are all lower than those of the cosponsorship and committee networks and higher than those of the alumni network, which are known for being too dense and too sparse, respectively.[69] The level of assortativity and clustering are also significantly lower than in the other networks, plausibly because the links are not generated by a shared activity or membership to observable political cliques. Betweenness, on the other hand, is higher than in the cosponsorship and alumni networks, reflecting an estimated higher level of political intermediation. Further evidence on the comparison between the estimated and observed networks can be found in Section A.5 in the online appendix.

[68]Another interesting feature is that most of the effective politicians tend to also have a higher number of connections, as shown in Panel (d).

[69]A common complaint with the cosponsorship network is that cosponsorships are affected by many factors that do not necessarily reflect social closeness between legislators. On the contrary, the alumni network uses only a limited source of information to draw social connections. See Battaglini and Patacchini [2019] for a discussion and survey of the literature.

Legislator characteristics. Table 1 shows how the legislators' individual characteristics directly affect their effectiveness (i.e. the β_is in ε_is). These are also reported in Table 3, where we show how the importance of these characteristics varies along the legislative *iter*. Perhaps unsurprisingly, the characteristics making legislators effective are different at differrent stages of the legislative process.

The estimated effects are in line with the existing theories of congressional politics (a literature that we discuss in Section A.6 in the online Appendix). Specifically, we find that committee chairs and leaders are associated with a higher probability of having bills that pass in the House. This is consistent with the idea that Congress members who are experienced and in positions of leadership are more effective. The same effect can be found for Congress members who have served in state legislatures with higher levels of professionalism. Interestingly, though, leaders and politicians who served in State legislatures with higher levels of professionalism appear less effective at earlier stages of the legislative process. This is consistent with the idea that the high level of effort required by legislators in certain institutional positions may result in a number of endorsed bills which is lower than that of the average House member (Volden and Wiseman [2014]). We detect nonlinear effects for seniority and the margin of victory. For seniority, we find a quadratic effect with increasing returns: higher seniority is therefore associated with greater effectiveness, but only after a few years of experience. This is in line with recent evidence on U.S. Senators which shows that freshmen have a lower effectiveness than the average Congress member (Volden and Wiseman [2018]). The impact of the margin of victory is concave, suggesting that electorally-safe members are more effective, but that the relative impact of electoral safety on legislative effectiveness exhibits decreasing returns. This is again fully in line with the evidence in Volden and Wiseman [2018], even though in their regression results with data on U.S. Senators, the estimates are not statistically significant. Non-white representatives, on the other hand, appear associated with a lower LES score, especially when looking at the proposal of bills. The delegation size has a positive impact on effectiveness at all stages of the legislative process, supporting the theory that legislators coming from larger congressional delegations may be more effective because they can find coalition partners among the members of their delegations. Politicians with more extreme ideologies are more likely to propose bills but much less likely to see them pass the House. Democrats were also significantly more likely to propose bills, but less likely to have them pass.

5.3 Counterfactual experiments

An advantage of our structural approach is that it allows us to perform a counterfactual analysis. That is, we can study how the social network changes as we change the fundamentals, and how these changes in turn affect the lawmakers' effectivenesses. We present here two exercises. First, we simulate the effects of a policy that moderates the ideologies of more extreme politicians and look at how the effects of such a policy impacts the relative effectiveness of different types of lawmakers. Second, we study the role of the alumni connections (i.e. "old boy" networks) in shaping social connections in the U.S. Congress. As discussed above, alumni ties play a significant role in determining lawmakers' social circles. We therefore ask the question: who would benefit from a policy that attenuates the importance of alumni ties in Congress?

Ideological extremism. To simulate the effects of a reduction in the ideologies of the most extreme lawmakers, we label as extremists the lawmakers with a DW-nominate score higher than some threshold t and consider a policy that reduces their scores to the average value of those below t. Such a reduction may be the result of a pledge of the two parties to put forward less ideological politicians or follow from campaign finance reform.[70] As discussed in the previous section, the DW-nominate score has a negative direct effect on lawmakers' effectiveness, so the policy has a direct positive effect on legislative productivity. We are interested in the spillover effects generated by this policy on the overall and relative change in productivity net of the direct policy effect.

To evaluate the policy, we simulate counterfactuals in which the ideologies of an increasing share of extreme legislators is moderated by setting t equal to 0.9, 0.8 and 0.7.[71] In Figure 8, we report the distribution of DW-nominate scores before the treatment (dashed line) and after the treatment (solid line) when the highest number of extreme politicians is moderated ($t = 0.7$). To clean out the mechanical effect of the treatment and focus only on the network-implied indirect effects, we consider the net effect of the policy change after subtracting the direct effect of the ideology change on the treated lawmakers.

Perhaps unsurprisingly given the positive spillovers, we find a positive effect of the reduction in extremism on the effectiveness of lawmakers net of the direct effect. The effect of the policy is to increase the number of bills passed in the House by roughly one percentage point. Perhaps more surprisingly, however, we find interesting "distributional" effects as well. These effects suggest that women, older lawmakers, and Republicans have more to benefit from the moderating policy, both in terms of total effectiveness and of successful bills. In Table 7, we show the effects as functions of demographic and political characteristics by reporting the results of regressions where the dependent variable is the change in productivity of the individual politicians following a moderating policy for the three different levels of t. Women and older lawmakers are more moderate on average, and so are less likely to be directly affected by the policy. However, they are relatively more connected to men and younger lawmakers, respectively, who are less moderate (on average) and thus more likely to be directly impacted by the policy (see Table A.4 in the online appendix). Replacing the more ideological elements in these groups, therefore, indirectly benefits women and older lawmakers. Republicans have higher values of the DW-nominate score, so they are also more likely to be affected by the policy. Our results, however, show that they tend to

[70] For example, Hall [2019] suggests that campaign finance reform could encourage more moderates to run by reducing the amount of time candidates spend fundraising and thus making it easier to run for office.

[71] The DW-nominate score ranges from 0 to 1. A score higher than 0.7 corresponds to roughly the 80th percentile of the distribution.

benefit from the policy even if we eliminate the direct effect of the reduction in ideology because they also have a higher number of connections within the party, which are with individuals more likely to be affected by the policy.

Alumni connections. As we have seen in the previous section, alumni connections reduce the cost of forming a social link between two lawmakers. It is natural to think that mitigating the importance of alumni ties may have important effects on the shape of the social network and, as a consequence, on the distribution of power in the U.S. Congress. For the most part, alumni ties were formed in the seventies and eighties, decades before the congresses we study, in periods in which "old boys" networks were more prevalent and tended to exclude women and other minorities. What if we eliminate these "old boys networks?" To study the effect of alumni connections in shaping social networks in Congress, we answer this question by simulating a scenario in which the alumni networks do not exist.[72]

One way to measure how the distribution of power changes is to look at how lawmakers' network centralities are affected by the policy. Table 8 shows the average change before and after the policy for five standard centrality measures and different groups of legislators.[73] The table shows that the policy has important redistributive effects. It increases the centrality of lawmakers from minority groups: for example, non-white politicians obtain higher closeness, degree, and clustering centrality. Women also gain higher centralities than men: they obtain a significantly higher eigenvector centrality, which means that they can get impulses from nodes eventually far from their direct contacts. As mentioned, these effects may be explained by the fact that "old boy" connections tend to discriminate against women and other minority groups, so removing these influences is beneficial for them.

Two other interesting effects emerge from Table 8. First, after eliminating the influence of the alumni connections, more senior legislators, chairs of committees, and Democrats gain centrality. This could be explained by the fact that links based on alumni ties partly crowd out links with legislators with higher experience, influence through committees, and in the majority. Legislators have a given budget of time to socialize, so favoring one type of link comes at the cost of other potentially useful links.[74] Second, legislators with more extreme ideologies and those with a low margin of victory in their election have lower centralities without their alumni connections because they cannot enjoy pre-existing connections and are less appealing links for others.

[72] Specifically, we set the positive coefficient estimated in Table 2 equal to zero (thus having the same cost in $\theta_{i,j}$ for a pair of legislators that attended the same school and a pair that did not) and substituting the censored links with random links to keep the network density constant.

[73] The baseline network before the policy is the estimated network.

[74] In the 109th Congress, Republicans controlled the Presidency and the majority in both the Senate and the House, but for the remainder of the sample, Democrats had a significant lead. In the 110th and 111th Congress, Republicans lost their majority both in the Senate and the House. In the 112th and 113th Congress, Democrats gained the Presidency and the majority in the Senate while Republicans held the House.

6 Other discussions and extensions

6.1 Two sided links

In the previous analysis, we assumed that links are one-sided: legislator i controlled $g_{i,j}$ (that is, a link that allows i to benefit from j) directly, at a cost $\frac{\lambda}{(1+\lambda)}\left(\frac{g_{i,j}}{\theta_{i,j}}\right)^{1+\frac{1}{\lambda}}$. In this section, we extend the analysis to allow for the possibility that links are two-sided. We now assume that a link i,j depends not only on the effort exerted by i versus j, denoted $\xi_{i,j}$, but also on the effort exerted by j versus i, denoted $\xi_{j,i}$. The cost of effort remains $\frac{\lambda}{(1+\lambda)}\left(\frac{\xi_{j,i}}{\theta_{i,j}}\right)^{1+\frac{1}{\lambda}}$, the link is now:

$$g_{i,j} = \xi_{i,j}^{\vartheta} \cdot \xi_{j,i}^{1-\vartheta} \tag{20}$$

If $\vartheta = 1$, then we are in the previous case, and if $\vartheta < 1$, then a link is the result of effort by both i and j: if j chooses $\xi_{i,j} = 0$, for example, than i cannot establish a link with j.

Under (20), i's problem finds $\xi^{\mathbf{i}}$ solving:

$$\max_{\xi^{\mathbf{i}}}\left\{\sum\nolimits_{j=1}^{n}\left[\alpha\delta\cdot\xi_{i,j}^{\vartheta}\cdot\xi_{j,i}^{1-\vartheta}\cdot E_j(G,\varepsilon) - \frac{\lambda}{(1+\lambda)}\left(\frac{\xi_{i,j}}{\theta_{i,j}}\right)^{1+\frac{1}{\lambda}}\right]\right\}. \tag{21}$$

From i and j's first order conditions we have:

$$\alpha\vartheta\delta\left(\frac{\xi_{j,i}}{\xi_{i,j}}\right)^{1-\vartheta}\theta_{i,j}^{\frac{1+\lambda}{\lambda}}E_j = \xi_{i,j}^{\frac{1}{\lambda}},\ \alpha\vartheta\delta\left(\frac{\xi_{i,j}}{\xi_{j,i}}\right)^{1-\vartheta}\theta_{j,i}^{\frac{1+\lambda}{\lambda}}E_i = \xi_{j,i}^{\frac{1}{\lambda}}$$

which gives us:

$$\frac{\xi_{i,j}}{\xi_{j,i}} = \left(\frac{E_j}{E_i}\right)^{\frac{1}{2(1-\vartheta)+1/\lambda}} \tag{22}$$

$$\xi_{i,j} = (\alpha\vartheta\delta)^{\lambda}\,\theta_{i,j}^{1+\lambda}\,(E_j)^{\lambda}\left(\frac{E_i}{E_j}\right)^{\frac{(1-\vartheta)\lambda}{2(1-\vartheta)+1/\lambda}},\ \xi_{j,i} = (\alpha\vartheta\delta)^{\lambda}\,\theta_{i,j}^{1+\lambda}\,(E_i)^{\lambda}\left(\frac{E_j}{E_i}\right)^{\frac{(1-\vartheta)\lambda}{2(1-\vartheta)+1/\lambda}}$$

And therefore:

$$g_{i,j} = (\alpha\vartheta\delta)^{\lambda}\,\theta_{i,j}^{1+\lambda}\cdot\left[(E_j)^{\lambda}\left(\frac{E_i}{E_j}\right)^{\frac{(1-\vartheta)\lambda}{2(1-\vartheta)+1/\lambda}}\right]^{\vartheta}\left[(E_i)^{\lambda}\left(\frac{E_j}{E_i}\right)^{\frac{(1-\vartheta)\lambda}{2(1-\vartheta)+1/\lambda}}\right]^{1-\vartheta} \tag{23}$$

Combining the solution of (20) with (7), we have:

Proposition 4. *In an interior solution with $g_{i,j} < \overline{g}$, a NCE is characterized by a vector E^* and* *a matrix G^* where E^* solve the system:*

$$E_i^* = \delta \cdot \sum\nolimits_{l \in \mathcal{N}} \left((\alpha \vartheta \delta)^{\lambda} \theta_{i,j}^{1+\lambda} \cdot \left[(E_l^*)^{\lambda} \left(\frac{E_i^*}{E_l^*} \right)^{\frac{(1-\vartheta)\lambda}{2(1-\vartheta)+1/\lambda}} \right]^{\vartheta} \left[(E_i^*)^{\lambda} \left(\frac{E_l}{E_i} \right)^{\frac{(1-\vartheta)\lambda}{2(1-\vartheta)+1/\lambda}} \right]^{1-\vartheta} E_l^* \right) + \varepsilon_i \quad (24)$$

for $i, j \in N$ and $G^* = (g_{i,j}^*)_{i,j \in \mathcal{N}}$ is given by (23).

Note that for $\vartheta = 1$, this expression is identical to (12). An advantage of (24) is that it allows us to estimate to what extent two-sided links are important in the data.

6.2 Alternative functional forms

For the empirical analysis presented above, we needed to assume some specific functional forms for the "production function" of effectiveness and the cost functions. From a theoretical point of view, most of our choices are not strictly necessary, in the sense that changing them would not prevent the empirical estimation. It is, however, clear that the details of the analysis depend on them. To illustrate this point, we consider a variant of the previous model in which:

$$E_i = \rho A_i \cdot (s_i)^{\alpha} (l_i)^{1-\alpha} + \varepsilon_i \quad (25)$$

Under (25), legislator i's characteristics not only affect E_i additively, through ε_i, they also affect it multiplicatively through A_i. In practice, in (25) we are allowing the spillover effect ρ to depend on the legislators' characteristics, i.e. $\rho_i = \rho \cdot A_i$.

Following the same steps as in Section 3, where we incorporate the optimal level of effort in (25), we obtain that in stage 2, when G is given, the effectiveness is given by:[75]

$$\mathbf{E} = (I - \delta \cdot \Gamma(A, \alpha) \cdot G)^{-1} \varepsilon \quad (26)$$

where $\Gamma(A, \alpha)$ is a diagonal matrix with ith diagonal element equal to $A_i^{\frac{1}{\alpha}}$. Condition (26) generalizes (7) by avoiding the need to impose that the social spillover is the same for all legislators: some may benefit more or less than others. In (26), the social spillover depends on the legislators characteristics affecting A_i. This generalization gives us additional flexibility when fitting the model.

Similarly as we did in Section 4.1, we now can specify $A_i = (\mathbf{X}_{i,r})' \varsigma_{\mathbf{1}} + \varsigma_2$, where $\varsigma_{\mathbf{1}}$ is a vector of coefficients to be estimated and ς_2 is a fixed effect.[76] Solving for the optimal social connections

[75] In (26) we are assuming that the linking technology is one sided as described in Section 2 for simplicity. There is no problem in generalizing the result using the "two sided links" technology described in Section 6.1.

[76] Naturally, we can also include random terms, such as ε_i in Section 4.1.

in the first stage as in Section 3.3, in an interior solution we have:

$$E_i = \delta \cdot [\alpha\delta]^{\lambda} \left[(\mathbf{X}_{i,r})'\varsigma_1 + \varsigma_2\right]^{\frac{1+\lambda}{\alpha}} \cdot \sum\nolimits_j [\theta_{i,j} E_j]^{1+\lambda} + \varepsilon_i \tag{27}$$

with $g_{i,j} = \left[\alpha\delta \cdot A_i^{\frac{1}{\alpha}}\right]^{\lambda} \cdot \theta_{i,j}^{1+\lambda} E_j^{\lambda}$. We can then use (27) in step C2 of our algorithm exactly as described in Section 4.2.

6.3 Unobserved factors

In our model, we deal with the endogeneity of the network by explicitly modeling its formation. Still, there may be correlated unobservable factors affecting both the cost of forming a link $\theta_{i,j,r}$ and the effectiveness $E_{i,r}$. Suppose that an unobserved characteristic of node i in network r, $\eta_{i,r}$, matters for the network formation process. Then we can extend model (19) in the following way:[77]

$$\chi_{i,j,r} = \iota + \gamma h_{i,j,r} + \sum\nolimits_l g(X_{i,l,r}, X_{j,l,r})\psi_l + m(\eta_{i,r}, \eta_{j,r})\kappa, \tag{28}$$

where $m(\cdot,\cdot)$ is a distance function and κ is a parameter. We allow the outcome error term $\epsilon_{i,r}$ to be correlated with $\eta_{i,r}$, with $\epsilon_{i,r} = \sum_{l=1}^{L} \mu_l \eta_{i,r}^l + u_{i,r}$.[78] In any step of the ABC algorithm (described in Section A.3), $\epsilon_{i,r}$ is thus generated as a function of $\eta_{i,r}$ and $\mu = (\mu_1, ..., \mu_L)$, which are estimated as the rest of the model's parameters.

The results associated with this extension of the empirical model are shown in column 2 of Table 2 for the network formation model (equation (28)) and in Table A.3 in the online appendix for the outcome equation (equation (17) augmented with $\epsilon_{i,r} = \sum_{l=1}^{L} \mu_l \eta_{i,r}^l + u_{i,r}$).[79] The results in Table 2 show that unobservable factors are important in shaping links (and thus social spillovers). The estimated coefficient is negative (as expected), indicating that similarity in unobserved characteristics is positively associated with the probability of forming a link. The important message in Table A.3 is that the evidence on the positive and statistically significant estimates of λ, and φ holds true.

By allowing for unobserved variables that are correlated with the legislators' effectiveness to affect the choice of social links, our approach is smiliar to Hsieh and Lee [2014] and Goldsmith-Pinkham and Imbens [2013]. Our analysis, however, differs in two important ways. First, in our model the true network is unobservable: only variables potentially determining the network

[77]The model can be further extended to a vector of unobservable characteristics.

[78]We define $m(\eta_i, \eta_j) = I(\eta_i \geq \sigma_\eta)I(\eta_j \geq \sigma_\eta)$, where $I(.)$ is an indicator function and σ_η is a threshold that we set equal to the variance of η. When the threshold is crossed, $m(.)$ switches to one. We set $L = 5$. $u_{i,r}$ is extracted from a normal distribution with zero mean and variance equal to σ_u. Other functional forms can be used. For example, we could assume $m(\eta_{r,i}, \eta_{r,j}) = \eta_{r,i} + \eta_{r,j}$ to capture degree heterogeneity (see Graham [2017]); or $m(\eta_{r,i}, \eta_{r,j}) = \eta_{r,i}\eta_{r,j}$ to include multiplicative effects. We experimented with other functional forms, distance definitions, thresholds, and L. Results remain qualitatively unchanged.

[79]In the first column of Table A.3, we also report our baseline estimates of Column (3) in Table 4 for comparison.

formation are observable. Second, and more importantly, the network endogeneity is explicitly micro-founded in our model. This is what gives us the nonlinear representation in (10).

6.4 Contextual effects

In equation (17) we have assumed that the idiosyncratic factor affecting i's effectiveness has the simple linear structure: $\varepsilon_r = \mathbf{X}_r\beta + \zeta_r + \epsilon_r$, where $\mathbf{X}_r$ is the $n \times L$ matrix of characteristics in Congress r. In this specification, i's effectiveness is not directly affected by the characteristics of other lawmakers, the so-called contextual effects. The characteristics of the other lawmakers affect the effectiveness of i indirectly because they affect the cost of a link through $\theta_{i,j}$ in our model (see equation (19)). An alternative specification is to allow for direct contextual effects from the other relevant lawmakers. This extension may allow us to better capture unobserved effects correlated among the lawmakers that may affect their effectiveness. We can include these exogenous controls in ε_r by adding the term HX, where H is the matrix of alumni connections. We use H to introduce contextual effects here because we can take this matrix as exogenous. In this case, ε_r becomes: $\varepsilon_r = \mathbf{X}_r\beta_{\mathbf{1}} + H_r\mathbf{X}_r\beta_2 + \zeta_r + \epsilon_r$. Table A.5 in the online appendix presents the results of our main regression when we add this term. As it can be seen, the analysis is qualitatively unchanged, with very small changes on the median value of the key parameters and their statistical significance. Most of the contextual effects are not significantly different from zero, with two exceptions: delegation size, which has a positive effect; and age, which has a negative effect. The results in Table A.6 in the online appendix show that the inclusion of contextual effects does not alter the estimates of coefficients in equation (19).

6.5 Other extensions: allowing for atomistic players, dynamic networks, and negative spillovers

To save space, we discuss three other extensions in the online appendix A.1. First, we show that the theoretical model can be extended to allow for both non-atomistic players who are "price takers" with respect to the effectiveness of the other players as in the NCE (the "followers"), and atomistic players who are not price takers ("the leaders"). This distinction allows us to reduce the dimensionality of the problem, while still permitting strategic behavior by the leaders. Second, we discuss how the NCE can be used as a building block of a more general dynamic theory of network formation, in which the social network at time t is a state variable in the network formation process at $t + 1$, a theory that we plan to develop in future research. Finally, in the third extension, we show how the basic model can be extended to allow for negative social spillovers among the players.

7 Conclusions

We have presented a model of the endogenous formation of social connections. In the model, agents first invest in their social connections with others, then exert effort, taking the connections as given. As an application, we use this model to study the social network of U.S. Congress members and their legislative effectiveness. The ability of agents to increase their legislative effectiveness depends on the social connections that they have previously established, the legislative effectiveness of the others with whom they are socially connected, their efforts, and their characteristics.

Two methodological contributions are at the core of our analysis. First, we introduce a new equilibrium concept that we call the Network Competitive Equilibrium. Drawing an analogy with the approach in competitive analysis to study "price taking" consumers, we assume that legislators take as given the expected effectiveness of other legislators when investing in their social connections. As prices in a competitive equilibrium, however, the vector of effectiveness needs to satisfy equilibrium conditions, i.e. to be consistent with the individual optimizing behavior of the legislators. The second contribution consists of the estimation technique, which is based on our characterization of the equilibrium and on the approximate Bayesian computation (ABC) method, an approach to draw Bayesian inference without an analytic likelihood function.

Using data from the 109th to 113th U.S. Congresses, we structurally estimate the model to gain insight on the legislators' social network and its effect on the legislators' effectiveness. The approach used in this work can more broadly be applied to study social networks in environments where social connections are endogenous and unobservable, and only a vector of outcomes (in our application the vector of legislators' effectiveness) is available for the estimation.

8 Appendix

8.1 Proof of Proposition 1

Let $g_{i,l}$ and $g_{i,j}$ be the links chosen by agent i with agents l and j. Agent i chooses the links solving (9) under the constraint that $g_{i,j} \in [0, \overline{g}]$ for all $j \in \mathcal{N}$. It is easy to see that, given $(E_j(G,\varepsilon))_{j\in\mathcal{N}}$, problem (9) is concave in $\mathbf{g_i} = (g_{i,1}, ..., g_{i,n})$ and, therefore, it has a unique solution whose necessary and sufficient condition is: $g^*_{i,j} \leq (\theta_{i,j})^{1+\lambda} \left(\alpha\delta E^*_j\right)^{\lambda}$, satisfied as an equality in an interior solution. Combining this with the necessary and sufficient condition (7) at $t = 2$, we have that an equilibrium $(G^*, \mathbf{E}^*)$ is characterized by the system in $n \cdot n$ equations in $n \cdot n$ variables (the n variables E^*_js and the $n(n-1)$ variables $g^*_{i,j}$ for $i \in N$ and $j \in N\backslash\{i\}$) described in (10) and (11) (given that the elements in the diagonal of G^* are all zero).[80] ■

8.2 Proof of Proposition 2

Define the mapping $T(\mathbf{x})$ from $[0,1]^n$ to $[0,1]^n$ as $T_i(\mathbf{x}) = \alpha^{\lambda} (\delta)^{1+\lambda} \sum_{l\in\mathcal{N}} (\theta_{i,l}x_l)^{1+\lambda} + \varepsilon_i$ for $i = 1, ..., n$. For any two $\mathbf{x}$, $\mathbf{y} \in [0,1]^n$, we can write:

$$\begin{aligned}
\|T(\mathbf{x}) - T(\mathbf{y})\| &= \alpha^{\lambda} \sum_i \left|\sum\nolimits_{l\in\mathcal{N}} (\delta\theta_{i,l})^{1+\lambda} \left[(x_l)^{1+\lambda} - (y_l)^{1+\lambda}\right]\right| \\
&\leq \alpha^{\lambda} \left(\delta\overline{\theta}\right)^{1+\lambda} \sum_i \left|\sum\nolimits_{l\in\mathcal{N}} 1_{\theta_{i,l}>0} \cdot \left[(x_l)^{1+\lambda} - (y_l)^{1+\lambda}\right]\right| \\
&\leq \alpha^{\lambda} \left(\delta\overline{\theta}\right)^{1+\lambda} \sum_i \sum\nolimits_{l\in\mathcal{N}} 1_{\theta_{i,l}>0} \cdot \left|(x_l)^{1+\lambda} - (y_l)^{1+\lambda}\right| \\
&= \alpha^{\lambda} \left(\delta\overline{\theta}\right)^{1+\lambda} \sum\nolimits_{l\in\mathcal{N}} \sum\nolimits_i 1_{\theta_{i,l}>0} \cdot \left|(x_l)^{1+\lambda} - (y_l)^{1+\lambda}\right| \\
&\leq \alpha^{\lambda}\overline{m} \left(\delta\overline{\theta}\right)^{1+\lambda} \sum\nolimits_{l\in\mathcal{N}} \left|(x_l)^{1+\lambda} - (y_l)^{1+\lambda}\right|
\end{aligned}$$

where $\|\mathbf{a} - \mathbf{b}\| = \sum_l |a_l - b_l|$ for any $\mathbf{a}, \mathbf{b} \in [0,1]^n$. The first inequality follows from the fact that $\theta_{i,l}$ is either zero or lower than $\overline{\theta}$, the second inequality follows from the triangle inequality (i.e. $|\sum x_l| \leq \sum |x_l|$); the second equality from the fact that we can invert the order of summation, and the last inequality from the fact that each agent l has at most $\overline{m}$ compatible j legislators who can establish a link. We can now write:

$$\begin{aligned}
\|T(\mathbf{x}) - T(\mathbf{y})\| &\leq \alpha^{\lambda}\overline{m} \left(\delta\overline{\theta}\right)^{1+\lambda} \sum\nolimits_{l\in\mathcal{N}} \left|(x_l)^{1+\lambda} - (y_l)^{1+\lambda}\right| \\
&\leq (1+\lambda)\, \alpha^{\lambda}\overline{m} \left(\delta\overline{\theta}\right)^{1+\lambda} \cdot \|\mathbf{x} - \mathbf{y}\|
\end{aligned}$$

[80]Since we are assuming that the diagonal values $\theta_{i,i}$ $\forall i$ are zeros, the values $g^*_{i,i}$ are also all zero. This implies that, effectively, the number of free variables is $n \cdot n$: the $n(n-1)$ values of $g^*_{i,j}$ for $i \neq j$ and the n values E^*_j for $i = 1, ..n$.

where the second inequality follows from the fact that for any $x_l, y_l \in [0,1]$, we have:

$$\begin{aligned} \left|x_l^{1+\lambda} - y_l^{1+\lambda}\right| &= \left|\int_y^x (1+\lambda)\, t^\lambda dt\right| \leq (1+\lambda) \left|\int_y^x dt\right| \\ &= (1+\lambda)\,|x_l - y_l| \end{aligned}$$

It follows that $T(x)$ is a contraction with a unique fixed-point if $\delta < \frac{1}{\theta}\left[1/\left((1+\lambda)\,\alpha^\lambda \overline{m}\right)\right]^{1/(1+\lambda)}$. ■

8.3 Proof of the result in Example 2

See online appendix. ■

8.4 Proof of Proposition 3

The stationary distribution associated with Algorithm C is defined by the fixed-point:

$$F_\nu^*(\omega) = \int_x Q_\nu\,(x \to \omega)\, F_\nu^*(x) dx \tag{29}$$

where $Q_\nu\,(x \to \omega)$ is the transition probability of Algorithm C. We need to prove that $f\,(\omega\,|\mathcal{D}_\nu\,)$ is the unique fixed-point of (29). Let ω be the vector of parameters to estimate and let $\mathcal{D}_\nu$ be defined as in Section 4.2. We have:

$$Q_\nu\,(\omega \to \omega') = q\,(\omega \to \omega')\,\Pr(\mathcal{D}_\nu\,|\omega'\,) \cdot h(\omega, \omega')$$

There are now two cases to consider. Assume first that:

$$\frac{\pi\,(\omega')\, q(\omega' \to \omega)}{\pi\,(\omega)\, q(\omega \to \omega')} \leq 1 \tag{30}$$

We can then write for any $\omega \neq \omega'$:

$$\begin{aligned} f\,(\omega\,|\mathcal{D}_\nu\,)\, Q_\nu\,(\omega \to \omega') &= f\,(\omega\,|\mathcal{D}_\nu\,) \cdot q\,(\omega \to \omega') \cdot \Pr(\mathcal{D}_\nu\,|\omega'\,) \cdot \frac{\pi\,(\omega')\, q(\omega' \to \omega)}{\pi\,(\omega)\, q(\omega \to \omega')} \\ &= \frac{\Pr(\mathcal{D}_\nu\,|\omega)\pi(\omega)}{\Pr\,(\mathcal{D}_\nu)} \cdot q\,(\omega' \to \omega) \cdot \Pr(\mathcal{D}_\nu\,|\omega'\,) \cdot \frac{\pi\,(\omega')}{\pi\,(\omega)} \\ &= \frac{\Pr(\mathcal{D}_\nu\,|\omega'\,)\pi(\omega')}{\Pr\,(\mathcal{D}_\nu)} \cdot q\,(\omega' \to \omega) \cdot \Pr(\mathcal{D}_\nu\,|\omega\,) \\ &= f\,(\omega'\,|\mathcal{D}_\nu\,) \cdot q\,(\omega' \to \omega) \cdot \Pr(\mathcal{D}_\nu\,|\omega\,) \\ &= f\,(\omega'\,|\mathcal{D}_\nu\,) \cdot Q_\nu\,(\omega' \to \omega) \end{aligned}$$

where in the last stage we used the fact that if (30) holds, then $\frac{\pi(\omega')q(\omega\to\omega')}{\pi(\omega)q(\omega'\to\omega)} > 1$ and thus $h(\omega', \omega) = 1$.

Assume now that (30) does not hold. We have for any $\omega \neq \omega'$:

$$\begin{aligned}
f(\omega|\mathcal{D}_\nu) Q_\nu(\omega \to \omega') &= f(\omega|\mathcal{D}_\nu) \cdot q(\omega \to \omega') \cdot \Pr(\mathcal{D}_\nu|\omega') \\
&= \frac{\Pr(\mathcal{D}_\nu|\omega)\pi(\omega)}{\Pr(\mathcal{D}_\nu)} \cdot q(\omega \to \omega') \cdot \Pr(\mathcal{D}_\nu|\omega') \\
&= \frac{\Pr(\mathcal{D}_\nu|\omega')\pi(\omega')}{\Pr(\mathcal{D}_\nu)} \cdot q(\omega' \to \omega) \cdot \Pr(\mathcal{D}_\nu|\omega) \frac{\pi(\omega)}{\pi(\omega')} \frac{q(\omega \to \omega')}{q(\omega' \to \omega)} \\
&= f(\omega'|\mathcal{D}_\nu) \cdot q(\omega' \to \omega) \cdot \Pr(\mathcal{D}_\nu|\omega) \cdot h(\omega', \omega) \\
&= f(\omega'|\mathcal{D}_\nu) \cdot Q_\nu(\omega' \to \omega)
\end{aligned}$$

From these two cases we conclude that:

$$f(\omega|\mathcal{D}_\nu) Q_\nu(\omega \to \omega') = f(\omega'|\mathcal{D}_\nu) \cdot Q_\nu(\omega' \to \omega) \tag{31}$$

If we integrate both sides of (31) by ω' we have:

$$\begin{aligned}
f(\omega|\mathcal{D}_\nu) &= f(\omega|\mathcal{D}_\nu) \int_x Q_\nu(\omega \to x)\, dx \\
&= \int_x Q_\nu(x \to \omega) f(x|\mathcal{D}_\nu)\, dx
\end{aligned}$$

which proves that $f(\omega|\mathcal{D}_\nu)$ is a fixed-point of (29) and a stationary distribution of the process. To see that $f(\omega|\mathcal{D}_\nu)$ is unique, note that $Q_\nu(x \to \omega)$ defines an irreducible Markov chain on $\mathcal{D}_\nu$ in which all states in $\mathcal{D}_\nu$ are recurrent and it thus admits a unique stationary distribution. ■

8.5 Setup of the Simulations in Section 4.3

We conduct Monte Carlo simulations generating the dependent variable from the following variant of model (17):

$$E_i = \alpha^\lambda \delta^{1+\lambda} \sum_{j \neq i} (\theta_{i,j} E_j)^{1+\lambda} + \beta X_i + \epsilon_i \tag{32}$$

Here X_i is a unidimensional random variable generated from a normal distribution $N(0, \sigma_x)$ and $\theta_{i,j}$ is generated from (18) where:

$$\chi_{i,j} = \upsilon(\iota + \gamma h_{i,j}). \tag{33}$$

For $h_{i,j}$ we consider the two frameworks described in Section 4.3. In the real world model, we set $h_{i,j}$ equal to the alumni network described in Section 5.1, in the Erdos-Renyi model we set p equal to the density of the alumni network. In Section 4.3, we also change p and show that our results are not sensitive to it. For both models, we set $\sigma_x = 1$, $\beta = 1$, $\alpha = 0.5$, $\lambda = 2$, $\sigma_\epsilon = 0.1$ and $\rho = 0.5$. The linking parameters are set as $\iota = -1$, $\gamma = 2$ and $\upsilon = 7$. The parameters $(\iota, \gamma, \psi_l, \alpha, \lambda, \rho)$ are chosen so that the resulting network matches the average values of the following moments in the cosponsorship networks of the Congresses in our sample (from the

109th to the 113th): the clustering of the politicians' cosponsorship network, the degree and the closeness of the committee network (in which two legislators are linked if they belong to the same committee) and the assortativity of the alumni. See panel (a) of Table A.1 and Table 6. All of the results discussed in Section 4.3 are robust to alternative configurations of the parameters.

References

[1] Acemoglu, D. and P. Azar (2020). Endogenous production networks. *Econometrica 88*(1), 33-82.

[2] Allen, F. and A. Babus (2009). Networks in finance. *The network challenge: strategy, profit, and risk in an interlinked world*. Wharton School Pub, 367-382.

[3] Allen, G. and D. Gale (2000). Financial contagion. *Journal of Political Economy 108*(1), 1-33.

[4] Angrist, J.D. and K. Lang (2004). Does school integration generate peer effects? Evidence from Boston's Metco Program. *American Economic Review 94*(5), 1613-1634.

[5] Badev, A (2017). Discrete Games in Endogenous Networks: Equilibria and Policy. Working paper.

[6] Ballester, C., A. Calvo-Armengol, and Y. Zenou (2006). Who's Who in networks. Wanted: The Key Player. *Econometrica 74*(5), 1403-1417.

[7] Battaglini, M., F.W. Crawford, E. Patacchini, and S. Peng (2020a). A Graphical Lasso Approach to Estimating Network Connections: The Case of U.S. Lawmakers. NBER Working paper.

[8] Battaglini, M., V. Leone Sciabolazza, and E. Patacchini (2020b). Effectiveness of Connected Legislators. *American Journal of Political Science 64*(4), 739-756

[9] Battaglini, M. and E. Patacchini (2018). Influencing Connected Legislators. *Journal of Political Economy 126*(6), 2277-2322.

[10] Battaglini, M. and E. Patacchini (2019). Social Networks in Policy Making. *Annual Review of Economics 11*, 473-494.

[11] Boucher, V. (2020). Equilibrium Homophily in Networks. *European Economic Review 123.*

[12] Boucher, V., Y. Bramoulle, H. Djebbari, and B. Fortin (2014). Do peers affect student achievement? Evidence from Canada using group size variation. *Journal of applied econometrics 29*(1), 91-109.

[13] Breza, E., A.G. Chandrasekhar, T.H. McCormick, and M. Pan (2017). Using Aggregated Relational Data to Feasibly Identify Network Structure Without Network Data. *American Economic Review 110*(8), 2454-2484.

[14] Cabrales, A., A. Calvo-Armengol, and Y. Zenou (2011). Social Interactions and spillovers. *Games and Economic Behavior 72*(2), 339-360.

[15] Calvo-Armengol, A., E. Patacchini, and Y. Zenou (2008). Peer Effects and Social Networks in Education. CEPR Discussion Paper 7060.

[16] Calvo-Armengol, A., E. Patacchini, and Y. Zenou (2009). Peer Effects and Social Networks in Education. *The Review of Economic Studies 76*(4), 1239-1267.

[17] Cameron, A. C., and P. K. Trivedi (2005). Microeconometrics: methods and applications. Cambridge university press.

[18] Canen, N., M.O. Jackson, and F. Trebbi (2020). Endogenous Networks and Legislative Activity. Working paper.

[19] Carvalho, V. M. (2014). From micro to macro via production networks. *Journal of Economic Perspectives 28*(4), 23-48.

[20] Carvalho, V. M. and A. Tahbaz-Salehi (2019). Production networks: A primer. *Annual Review of Economics* 11, 635-663.

[21] Christakis, N.A., J.H. Fowler, G.W. Imbens, and K. Kalyanaraman (2020). An Empirical Model for Strategic Network Formation. *The Econometric Analysis of Network Data, 123-148.*

[22] Cohen, L., A. Frazzini, and C.J. Malfoy (2008). The Small World of Investing: Board Connections and Mutual Fund Returns. *Jounal of Political Economy 116*(5), 951-979.

[23] Cohen, L. and C. Malloy (2014). Friends in High Places. *American Economic Journal: Economic Policy 6*(3), 63-91.

[24] De Paula, A., I. Rasul, and P. Souza (2018). Recovering Social Networks from Panel Data: Identification, Simulations, and an Application. Cemmap Working paper 17/18.

[25] De Paula, A., S. Richards-Shubik, and E. Tamer (2018b). Identifying Preferences in Networks with Bounded Degree. *Econometrica 86*(1), 263-288.

[26] Denbee E., C. Juliard, Y. Li, and K. Yuan (2020). Network Risk and Key Players: A Structural Analysis of Interbank Liquidity. *Journal of Financial Economics, forthcoming.*

[27] Drehmann, M. and N. A. Tarashev (2011). Systemic importance: some simple indicators. BIS Quarterly Review, March.

[28] Fiorina, M.P. (1978). Legislative Facilitation of Government Growth: Universalism and Reciprocity Practices in Majority Rule Institutions. California Institute of Technology, Social Science Working paper 226.

[29] Florens, J.P. and A. Simoni (2011). Bayesian Identification and Partial Identification. Working paper.

[30] Florens, J.P., M. Mouchart and J.M. Rolin (1990). *Elements of Bayesian Statistics*. Routledge.

[31] Fowler, J.H. (2006). Connecting the Congress: A study of Cosponsorship Networks. *Political Analysis 14*(4), 456-487.

[32] Francois, O. and G. Laval (2011). Deviance information criteria for model selection in approximate Bayesian computation. *Statistical Applications in Genetics and Molecular Biology 10*(1).

[33] Fruchterman, T.M. and E.M. Reingold (1991). Graph Drawing by Force-Directed Placement. *Software: Practice and Experience 21*(11), 1129-1164.

[34] Furfine, C. H. (1999). The microstructure of the federal funds market. *Financial Markets, Institutions & Instruments 8*(5), 24-44.

[35] Furfine, C. H. (2003). Interbank Exposures: Quantifying the Risk of Contagion. *Journal of Money, Credit and Banking 35*(1), 111-128.

[36] Goldsmith-Pinkham, P. and G.W. Imbens (2013). Social Networks and the Identification of Peer Effects. *Journal of Business and Economic Statistics 31*(3), 253-264.

[37] Graham, B. S. (2017). An Econometric Model of Network Formation with Degree Heterogeneity. *Econometrica 85*(4), 1033-1063.

[38] Granovetter, M.S. (1973). The Strength of Weak Ties. *American Journal of Sociology 78*(6). 1360-1380.

[39] Hall, A. (2019). Want to Reduce Polarization in Congress? Make Moderates a Better Job Offer. Working paper.

[40] Hanushek, E. A., J.F. Kain, J.M. Markman, and S.G. Rivkin (2003). Does Peer Ability Affect Student Achievement?. *Journal of applied econometrics 18*(5), 527-544.

[41] Hastings, W.K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika 57*(1), 97-109.

[42] Hochberg, Y. V., A. Ljungqvist, and Y. Lu (2007). Whom You Know Matters: Venture Capital Networks and Investment Performance. *The Journal of Finance 62*(1), 251-301.

[43] Hsieh, C.S., and L.F. Lee (2014). A Social Interactions Model with Endogenous Friendship Formation and Selectivity. *Journal of Applied Econometrics 31*(2), 301-319.

[44] Jackson, M., and B. Rogers (2007). Meeting Strangers and Friends of Friends: How Random Are Social Networks? *American Economic Review 97*(3), 890-915.

[45] Jackson M. and A. Wolinsky (1996). A Strategic Model of Social and Economic Networks. *Journal of Economic Theory 71*(1), 44-74.

[46] Jeydel, A. and A.J. Taylor (2003). Are Women Legislators Less Effective? Evidence from the U.S. House in the 103rd-105th Congress. *Political Research Quarterly 56*(1), 19-27.

[47] Kang, C. (2007). Classroom Peer Effects and Academic Achievement: Quasi-randomization Evidence from South Korea. *Journal of Urban Economics 61*(3), 458-495.

[48] Kass, R.E. and L. Wasserman (1996). The Selection of Prior Distributions by Formal Rules. *Journal of the American Statistical Association 91*(435), 1343-1370.

[49] Kirkland, J.H. (2011). The Relational Determinants of Legislative Success: Strong and Weak Ties Between Legislators. *Journal of Politics 73*(3), 887-898.

[50] Konig, M.D. (2016). The Formation of Networks with Local Spillovers and Limited Observability. *Theoretical Economics 11*(3), 813-863.

[51] Krehbiel, K. (1992). *Information and Legislative Organization*. University of Michigan Press.

[52] Lancaster, T. (2004). *An Introduction to Modern Bayesian Econometrics*. Oxford Blackwell.

[53] Lindley, D.V. (1971). *Bayesian Statistics: a Review*. Philadelphia, SIAM.

[54] Lee, L., E. Patacchini, Y. Zenou, and X. Liu (2012). Who is the Key Player? A Network Analysis of Juvenile Delinquency. *Journal of Business & Economic Statistics, forthcoming*.

[55] Manresa, E. (2016). Estimating the Structure of Social Interactions Using Panel Data. CEMFI Working paper.

[56] Marjoram, P., J. Molitor, V. Plagnol, and S. Tavare (2003). Markov Chain Monte Carlo Without Likelihoods. *Proceedings of the National Academy of Sciences 100*(26), 15324-15328.

[57] McCarty, N.M., K.T. Poole, and H. Rosenthal (1997). *Income Redistribution and the Realignment of American Politics*, AEI Press.

[58] Mele, A. (2017). A Structural Model of Dense Network Formation. *Econometrica 85*(3), 825-850.

[59] Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller (1953). Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics 21*(6), 1087-1092.

[60] Mistrulli , P. (2007). Assessing financial contagion in the interbank market: Maximum entropy versus observed interbank lending patterns. *Journal of Banking and Finance 35*(5), 1114-1127.

[61] Miyauchi, Y. (2016). Structural Estimation of Pairwise Stable Networks with Nonnegative Externality. *Journal of Econometrics 195*(2), 224-235.

[62] Newman, M. (2010). *Networks: An Introduction.* Oxford University Press.

[63] Oberfield, E. (2018). A Theory of Input-Output Architecture. *Econometrica 86*(2), 559-589.

[64] Padro I Miquel, G. and J.M. Snyder Jr. (2006). Legislative Effectiveness and Legislative Careers. *Legislative Studies Quarterly 31*(3), 347-381.

[65] Patacchini, E. and Y. Zenou (2012). Juvenile Delinquency and Conformism. *The Journal of Law, Economics, & Organization 28*(1), 1-31.

[66] Peltonen T., M. Rancan and P. Sarlin (2015). Interconnectedness of the Banking Sector as a Vulnerability to Crises. *International Journal of Finance and Economics 24*(2)*, 963-990.*

[67] Peng, S. (2019). Heterogeneous Endogenous Effects in Networks. Working Paper.

[68] Rendon, S.R. (2013). Fixed and Random Effects in Classical and Bayesian Regression. *Oxford Bulletin of Economics and Statistics 75*(3), 460-476.

[69] Roberts, G.O., A. Gelman, and W.R. Gilks (1997). Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms. *The Annals of Applied Probability 7*(1), 110-120.

[70] Roberts, G.O., and J.S. Rosenthal (2001). Optimal Scaling for Various Metropolis-Hastings Algorithms. *Statistical Science 16*(4), 351-367.

[71] Roberts, G.O., and J.S. Rosenthal (2009). Examples of Adaptive MCMC. *Journal of Computational and Graphical Statistics 18*(2), 349-367.

[72] Rose, C. (2018). Identification of Spillover Effects Using Panel Data. Working Paper.

[73] Rubin, D.B. (1984). Bayesianly Justifiable and Relevant Frequency Calculations for the Applied Statistician. *The Annals of Statistics 12*(4), 1151-1172.

[74] Sheng, S. (2018). A Structural Econometric Analysis of Network Formation. *Econometrica 88*(5), 1829-1858.

[75] Smith, T.E., and J.P. LeSage (2004). A Bayesian Probit Model with Spatial Dependencies. *Spatial and Spatiotemporal Econometrics*, Emerald Group Publishing Limited, 127-160.

[76] Souza, P. (2014). Estimating Network Effects Without Network Data. Working paper.

[77] Spiegelhalter, D.J., N.G. Best, B.P. Carlin, and A. Van Der Linde (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B 64*(4), 1-34.

[78] Squire, P. (1992). Legislative Professionalization and Membership Diversity in State Legislatures. *Legislative Studies Quarterly 17*(1), 69-79.

[79] Upper, C. (2006). Contagion Due to Interbank Credit Exposures: What do we know, why do we know it, and what should we do now?. Working Paper, Bank for International Settlements.

[80] Victor, J.N., A.H. Montgomery, and M. Lubell (2016). *The Oxford Handbook of Political Networks*, Oxford University Press.

[81] Volden, C., and A.E. Wiseman (2011). Breaking Gridlock: The Determinants of Health Policy Change in Congress. *Journal of Health Politics, Policy and Law 36*(2), 227-264.

[82] Volden, C., and A.E. Wiseman (2014). *Legislative Effectiveness in the United States Congress: The Lawmakers*. Cambridge University Press.

[83] Volden, C., and A.E. Wiseman (2018). Legislative Effectiveness in the United States Senate. *Journal of Politics 80*(2), 731-735.

[84] Volden, C., A.E. Wiseman, and D.E. Wittmer (2013). When Are Women More Effective Lawmakers Than Men? *American Journal of Political Science 57*(2), 326-341.

[85] Weingast, B. (1979). A Rational Choice Perspective on Congressional Norms. *American Journal of Political Science 23*(2), 245-262.

[86] Weiss, G., and A. von Haeseler (1998). Inference of Population History Using a Likelihood Approach. *Genetics 149*(3), 1539-1546.

[87] Wiseman, A.E., and Wright, J.R. (2008). The Legislative Median and Partisan Policy. *Journal of Theoretical Politics 20*(1), 5-29.

Tables and Figures

Table 1: ESTIMATION RESULTS

Dependent variable: Legislative Effectiveness Score (LES)	
φ	0.0277 ***
	[1.0000]
λ	0.5980 ***
	[1.0000]
Party	-0.0124 ***
	[0.0000]
Gender	0.0012
	[0.7295]
Non white	-0.0042 ***
	[0.0000]
Seniority	-0.0001
	[0.4730]
Seniorty^2	0.0001 *
	[0.9489]
DW ideology	-0.0093 ***
	[0.0000]
Margin	0.0813 ***
	[1.0000]
Margin^2	-0.0493 ***
	[0.0000]
Committee chair	0.1393 ***
	[1.0000]
Powerful committee	-0.0083 **
	[1.0000]
Delegation size	0.0008 ***
	[0.9952]
Leader	-0.0026 *
	[0.0820]
State legislative experience	-0.0021
	[0.1765]
State legislative experience *	-0.0151 ***
State legislative professionalism	[0.0000]
Age	0.0002
	[0.8861]
State fixed effects	Yes
Topic fixed effects	Yes
Congress fixed effects	Yes
N. Obs.	2,176

NOTES. Estimates of parameters in equation (17). 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 2: LINK FORMATION

Dependent variable: probability of forming a link		
	(1)	(2)
Link in alumni network	0.2310 ***	0.1979 ***
	[1.0000]	[1.0000]
Seniority [1 = same quartile]	0.5206 ***	0.1571 ***
	[1.0000]	(1.0000)
Seniority i	0.0924 ***	0.1015 ***
	[1.0000]	(1.0000)
Seniority j	0.0246 ***	0.0291 ***
	[1.0000]	(1.0000)
Same state [1 = yes]	2.0048 ***	1.9954 ***
	[1.0000]	(1.0000)
Same topic [1 = yes]	0.2344 ***	0.2106 ***
	[1.0000]	(1.0000)
Leader [1 = both leaders]	-0.2899 ***	-0.2935 ***
	[0.0000]	(0.0000)
Same gender [1 = yes]	-0.5456 ***	-0.5181 ***
	[0.0000]	(0.0000)
Same race [1 = both white or both non white]	-0.0547 ***	-0.0802 ***
	[0.0000]	(0.0000)
Same party [1 = yes]	2.4994 ***	2.4857 ***
	[1.0000]	(1.0000)
Age [1 = same quartile]	0.1559 ***	0.1441 ***
	[1.0000]	(1.0000)
Unobservables [1 = both above the threshold]		0.0203 **
		(0.9500)
Intercept	-7.0805 ***	-7.1328 ***
	[0.0000]	[0.0000]
N. Obs.	2,176	2,176

NOTES. Estimates of parameters in equation (19) are reported in column (1), and estimates of parameters in equation (28) are reported in column (2). 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 legislator j respectively. The rest of the independent variables are dummies that capture 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 3: ESTIMATION RESULTS
-LEGISLATIVE EFFECTIVENESS DECOMPOSITION-

Dependent variable:	LES	BILL	AIC-ABC	PASS-LAW
φ	0.0277 ***	0.0269 ***	0.0218 ***	0.0207 ***
	[1.0000]	[1.0000]	[1.0000]	[1.0000]
λ	0.5980 ***	0.6293 ***	0.0164 ***	0.0031 ***
	[1.0000]	[1.0000]	[1.0000]	[1.0000]
Party	-0.0124 ***	0.0129 ***	-0.0321 ***	-0.0161 ***
	[0.0000]	[1.0000]	[0.0000]	[0.0000]
Gender	0.0012	0.0080 ***	-0.0080 ***	-0.0027 **
	[0.7295]	[1.0000]	[0.0000]	[0.0132]
Non white	-0.0042 ***	-0.0128 ***	0.0019	-0.0026
	[0.0000]	[0.0000]	[0.8322]	[0.1938]
Seniority	-0.0001	-0.0004	-0.0019 **	-0.0021 ***
	[0.4730]	[0.2669]	[0.0177]	[0.0000]
Seniorty^2	0.0001 *	0.0001 ***	0.0002 ***	0.0002 ***
	[0.9489]	[0.9901]	-10.000	[1.0000]
DW ideology	-0.0093 ***	0.0139 ***	-0.0314 ***	-0.0202 ***
	[0.0000]	[1.0000]	[0.0000]	[0.0000]
Margin	0.0813 ***	0.0923 ***	0.0659 ***	0.0190 ***
	[1.0000]	[1.0000]	[1.0000]	[1.0000]
Margin^2	-0.0493 ***	-0.0641 ***	-0.0385 ***	-0.0053 ***
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Committee chair	0.1393 ***	0.0611 ***	0.1888 ***	0.1315 ***
	[1.0000]	[1.0000]	[1.0000]	[1.0000]
Powerful committee	-0.0083 **	-0.0096 ***	-0.0107 ***	-0.0025 *
	[0.0417]	[0.0000]	[0.0000]	[0.0612]
Delegation size	0.0008 ***	0.0012 ***	0.0010 **	0.0005 **
	[0.9952]	[1.0000]	[0.9643]	[0.9610]
Leader	-0.0026 *	-0.0133 ***	0.0048 ***	0.0024 ***
	[0.0820]	[0.0000]	[0.9983]	[0.9909]
State legislative experience	-0.0021	-0.0020 *	0.0000	0.0012
	[0.1765]	[0.0960]	[0.5010]	[0.7326]
State legislative experience *	-0.0151 ***	-0.0291 ***	0.0062	0.0088 ***
State legislative professionalism	[0.0000]	[0.0000]	[0.8594]	[1.0000]
Age	0.0002	-0.0004	0.0001	0.0000
	[0.8861]	[0.2740]	[0.6672]	[0.6101]
State fixed effects	Yes	Yes	Yes	Yes
Topic fixed effects	Yes	Yes	Yes	Yes
Congress fixed effects	Yes	Yes	Yes	Yes
N. Obs.	2,176	2,176	2,176	2,176

NOTES. Estimates of parameters in equation (17). LES: legislative effectiveness. BILL: bills proposed in the House of Representatives. AIC-ABC: bills that received any action in committee and bills that received some action beyond the committee stage but have not passed in the House. PASS-LAW: bills that have passed in the House and those that have become law. 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 4: ESTIMATION RESULTS
-COMPARISON OF NESTED MODELS-

Dependent variable: Legislative Effectiveness Score (LES)	No network (1)	Exogenous network (2)	Endogenous network (3)
φ		0.0331 ***	0.0277 ***
		[1.0000]	[1.0000]
λ	-		0.5980 ***
			[1.0000]
Party	-0.0043	-0.0090 ***	-0.0124 ***
	[0.1877]	[0.0000]	[0.0000]
Gender	0.0021	-0.0001	0.0012
	[0.6886]	[0.4772]	[0.7295]
Non white	-0.0064 *	-0.0078 ***	-0.0042 ***
	[0.0964]	[0.0000]	[0.0000]
Seniority	0.0004	0.0001	-0.0001
	[0.6635]	[0.5374]	[0.4730]
Seniority^2	0.0001 **	0.0001 *	0.0001 *
	[0.9520]	[0.9341]	[0.9489]
DW ideology	-0.0115	-0.0093 ***	-0.0093 ***
	[0.1320]	[0.0000]	[0.0000]
Margin	0.0805 ***	0.0806 ***	0.0813 ***
	[0.9999]	[1.0000]	[1.0000]
Margin^2	-0.0505 ***	-0.0484 ***	-0.0493 ***
	[0.0067]	[0.0000]	[0.0000]
Committee chair	0.1421 ***	0.1427 ***	0.1393 ***
	[1.0000]	-10000,00	[1.0000]
Powerful committee	-0.0105 ***	-0.0107 ***	-0.0083 **
	[0.0021]	[0.0000]	[0.0417]
Delegation size	-0.0037 **	0.0011 ***	0.0008 ***
	[0.0249]	[1.0000]	[0.9952]
Leader	-0.0039	-0.0001	-0.0026 *
	[0.2865]	[0.4723]	[0.0820]
State legislative experience	-0.0019	-0.0001	-0.0021
	[0.3762]	[0.4927]	[0.1765]
State legislative experience *	-0.0130	-0.0134 ***	-0.0151 ***
State legislative professionalism	[0.2255]	[0.0000]	[0.0000]
Age	0.0003 *	0.0002	0.0002
	[0.9050]	[0.8936]	[0.8861]
State fixed effects	Yes	Yes	Yes
Topic fixed effects	Yes	Yes	Yes
Congress fixed effects	Yes	Yes	Yes
Partial F test		4850.73	584.70
p-value		0.0000	0.0000
pseudo- R_1^2	0.4392	0.9343	0.9492
pseudo- R_2^2	0.3868	0.7919	0.8209
MSE	0.5608	0.0164	0.0127
MASD	0.5615	0.0699	0.0692
DIC	2.2123	1.8425	1.8422
N. Obs.	2,176	2,176	2,176

NOTES. ABC estimates of parameters in equation (17). The median of the posterior distribution estimated with the ABC algorithm is reported for all specifications. The empirical p-value of zero on the estimated posterior is reported in brackets for each specification. R_1^2 and R_2^2 are pseudo R^2, computed as $1-(Q_{fit}-Q_0)/(Q_{min}-Q_0)$, where Q denotes the function to be minimized, Q_0 is its worst possible value, Q_{min} is its minimum possible value, Q_{fit} is the value estimated. We set $Q=\Sigma_{i=1}^N(y_i-\hat{y}_i)^2/(N-K)=MSE$, $Q_{min}=0$ and $Q_0=\{[max_i(y_i)-min_i(y_i)]/2\}^2$ for R_1^2 and $Q=\Sigma_{i=1}^N(|y_i-\hat{y}_i|/)N-K)$, $Q_{min}=0$ and $Q_0=[max_i(y_i)-min_i(y_i)]/2$ for R_2^2, where K is the number of parameters. See Cameron and Trivedi [2005], chapter 8.7. The MASD is the average absolute distance of 500 samples simulated using the estimated parameters from the real vector of outcomes. The DIC is computed following the procedure in Francois and Laval [2011], DIC_1. The partial F test in column (3) compares the unrestricted model with the one with $\lambda=0$, the partial F test in column (2) compares the latter with the one with $\varphi=\lambda=0$. 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 5: ESTIMATION RESULTS
-COMPARISON OF DIFFERENT MODELS AND METHODS-

Dependent variable: Legislative Effectiveness Score (LES)

	SAR (1)	Two Step I (2)	Two Step II (3)	Endogenous (4)
φ	0.5517 ***	0.3562 ***	0.3677 ***	0.0277 ***
	[1.0000]	[1.0000]	[1.0000]	-10000
λ	-	-	-	0.5980 ***
				-10000
Party	-0.0077 *	-0.0076 *	-0.0076 *	-0.0124 ***
	[0.0629]	[0.0646]	[0.0649]	[0.0000]
Gender	0.0029	0.0029	0.0029	0.0012
	[0.7512]	[0.7524]	[0.7529]	[0.7295]
Non white	-0.0049	-0.0049	-0.0049	-0.0042 ***
	[0.1627]	[0.1598]	[0.1648]	[0.0000]
Seniority	0.0002	0.0002	0.0002	-0.0001
	[0.5620]	[0.5615]	[0.5600]	[0.4730]
Seniority^2	0.0001 **	0.0001 **	0.0001 **	0.0001 *
	[0.9745]	[0.9758]	[0.9754]	[0.9489]
DW ideology	-0.0114	-0.0116	-0.0117	-0.0093 ***
	[0.1304]	[0.1285]	[0.1273]	[0.0000]
Margin	0.0843 ***	0.0844 ***	0.0844 ***	0.0813 ***
	[1.0000]	[1.0000]	[1.0000]	-10000
Margin^2	-0.0527 ***	-0.0523 ***	-0.0524 ***	-0.0493 ***
	[0.0050]	[0.0047]	[0.0048]	[0.0000]
Committee chair	0.1412 ***	0.1411 ***	0.1411 ***	0.1393 ***
	[1.0000]	[1.0000]	[1.0000]	-10000
Powerful committee	-0.0106 ***	-0.0105 ***	-0.0106 ***	-0.0083 **
	[0.0019]	[0.0016]	[0.0011]	[0.0417]
Delegation size	-0.0034 **	-0.0034 **	-0.0034 **	0.0008 ***
	[0.0330]	[0.0340]	[0.0344]	[0.9952]
Leader	-0.0031	-0.0030	-0.0031	-0.0026 *
	[0.3292]	[0.3340]	[0.3283]	[0.0820]
State legislative experience	-0.0012	-0.0011	-0.0011	-0.0021
	[0.4254]	[0.4305]	[0.4263]	[0.1765]
State legislative experience *	-0.0153	-0.0155	-0.0152	-0.0151 ***
State legislative professionalism	[0.1881]	[0.1865]	[0.1874]	[0.0000]
Age	0.0003 *	0.0003 *	0.0003 *	0.0002
	[0.9326]	[0.9311]	[0.9324]	[0.8861]
State fixed effects	Yes	Yes	Yes	Yes
Topic fixed effects	Yes	Yes	Yes	Yes
Congress fixed effects	Yes	Yes	Yes	Yes
pseudo- R_1^2	0.4453	0.2953	0.6194	0.9492
pseudo- R_2^2	0.5409	0.1618	0.7916	0.8209
MSE	0.5608	0.8382	0.2084	0.0127
MASD	0.5617	0.6633	1.3957	0.0692
DIC	2.1102	2.2310	2.6424	1.8422
N. Obs.	2,176	2,176	2,176	2,176

NOTES. ABC estimates of parameters in equation (17). The median of the posterior distribution estimated with the ABC algorithm is reported for all specifications. The empirical p-value of zero on the estimated posterior is reported in brackets for each specification. In column (1)-(3), $\lambda = 0$. In column (1), $\Theta = H$, the alumni network. In column (2), $\Theta = G^{Cosp}$ and $\sum_j \tilde{\epsilon}_{i,j}$ is added as a regressor in equation (17), where $\tilde{\epsilon}_{i,j} = g_{i,j}^{Cosp} - \hat{g}_{i,j}^{Cosp}$. See Battaglini et al. [2020] for more details. In column (3), $\Theta = \hat{G}^{Cosp}$, where $\hat{g}_{i,j}^{Cosp} = \hat{\gamma} h_{i,j} + \sum_l g(X_{i,l}, X_{j,l})\hat{\psi}_l$ and gender, age and race are included in the X. $g(.)$ is computed as described in Section 4.1. $g_{i,j}^{Cosp} = 1$ if i cosponsored at least 10 bills proposed by j, 0 otherwise. R_1^2 and R_2^2 MASD and DIC are computed as described in Table 4. 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 6: COMPARISON WITH OTHER NETWORKS - NETWORK-LEVEL STATISTICS -

	NETWORKS			
	ESTIMATED (1)	COSPONSORSHIP (2)	COMMITTEE (3)	ALUMNI (4)
Density	0.0037	0.1069	0.0401	0.0006
Assortativity	2.5028	20.3642	49.3329	6.7297
Closeness	0.0660	0.1533	0.1142	0.0027
Betwenness	0.0015	0.0004	0.0028	0.0011
Degree	0.0114	0.0904	0.0565	0.0067
Clustering	0.0565	0.6561	0.6107	0.5853

NOTES. 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 network-level statistics are compared on a pooled network of five Congresses. See Newman [2010] for the definition of network-level statistics. The alumni network is defined in Section 5.1. In the cosponsorship network, the ij_{th} element is 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.

Table 7: COUNTERFACTUAL ANALYSIS
- IDEOLOGICAL EXTREMISM -

DW ideology threshold	$t = 0.9$		$t = 0.8$		$t = 0.7$	
Dependent variable: change in	LES	PASS-LAW	LES	PASS-LAW	LES	PASS-LAW
Party	-0.0001 ***	-0.0001 ***	-0.0001 ***	-0.0002 ***	-0.0001 ***	-0.0001 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Gender	0.0000	0.0000	0.0001 **	0.0001 **	0.0002 ***	0.0001 **
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Non white	0.0000	0.0000	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Seniority	0.0000	-0.0000	0.0000	0.0000	0.0000 ***	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Seniority^2	0.0000	0.0000 ***	-0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Margin	0.0000	0.0000	-0.0000 *	0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Margin^2	-0.0000	-0.0000	0.0000	-0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Committee chairs	0.0001 **	0.0000 *	0.0000	0.0000 **	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Powerful committee	-0.0000	0.0000	0.0000	0.0000	0.0000 *	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Delegation size	-0.0000	0.0000	0.0000	-0.0001 ***	0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Leader	-0.0000	0.0000	0.0000	0.0000 **	0.0000 *	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
State legislative experience	-0.0000	-0.0000	0.0000	-0.0000	0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
State legislative experience*	0.0000	0.0000	0.0000	0.0000	-0.0000	-0.0000
State legislative professionalism	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Age	0.0001 **	0.0001 **	0.0001 **	0.0002 ***	0.0001 **	0.0001 **
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
N. Obs.	2,176	2,176	2,176	2,176	2,176	2,176

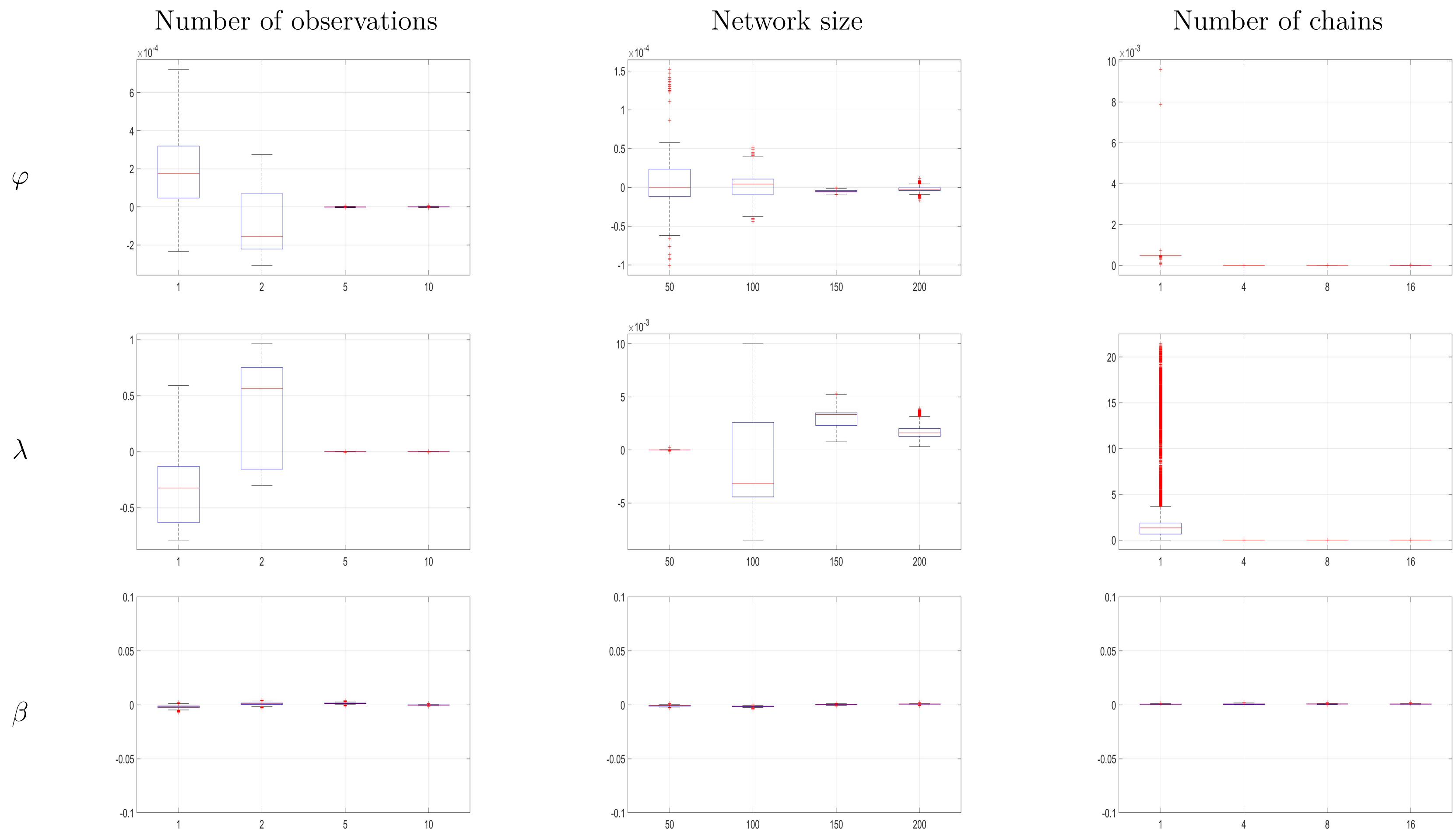
NOTES. OLS estimates. Standard errors are in parentheses. The dependent variable is the change in LES or PASS-LAW. LES: legislative effectiveness as defined in Section 5. PASS-LAW: bills that have passed in the House and those that have become law. The threshold t is the value of DW ideology after which the values are replaced with the mean of observations below t. 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 8: COUNTERFACTUAL ANALYSIS
- ALUMNI CONNECTIONS -

Centralities:	Eigenvector	Betweenness	Closeness	Degree	Clustering
Female	0.0024	0.0000	5.4379	0.0077	0.0042
T test p-value	*0.0330*	*0.9396*	*0.0010*	*0.0012*	*0.1196*
Non white	-0.0007	-0.0001	5.2283	0.0079	0.0030
T test p-value	*0.1362*	*0.0105*	*0.0268*	*0.0033*	*0.2596*
Chair	0.0047	0.0000	4.3472	0.0044	-0.0028
T test p-value	*0.0208*	*0.9970*	*0.7091*	*0.1395*	*0.9992*
Democrat	0.0020	0.0001	5.8566	0.0085	0.0013
T test p-value	*0.0003*	*0.0004*	*0.0000*	*0.0000*	*0.0436*
Age > sample mean	0.0011	-0.0000	4.8180	0.0053	-0.0022
T test p-value	*0.2935*	*0.1192*	*0.0321*	*0.0066*	*0.7591*
Seniority > sample mean	0.0013	-0.0000	4.5733	0.0059	-0.0007
T test p-value	*0.1319*	*0.7404*	*0.8963*	*0.0063*	*0.3669*
DW ideology > sample mean	-0.0005	0.0000	3.7319	0.0012	-0.0070
T test p-value	*0.0016*	*0.0058*	*0.0000*	*0.0000*	*0.0417*
Margin of victory < sample mean	-0.0002	0.0000	4.2398	0.0026	-0.0055
T test p-value	*0.0188*	*0.0181*	*0.0096*	*0.0016*	*0.1832*
Delegation size > sample mean	0.0011	-0.0000	4.5881	0.0040	0.0004
T test p-value	*0.2843*	*0.0359*	*0.7961*	*0.5964*	*0.1346*
Leg. experience > sample mean	0.0009	0.0000	4.6154	0.0037	-0.0024
T test p-value	*0.4892*	*0.9641*	*0.6120*	*0.3887*	*0.8418*
N. Obs.	2,176	2,176	2,176	2,176	2,176

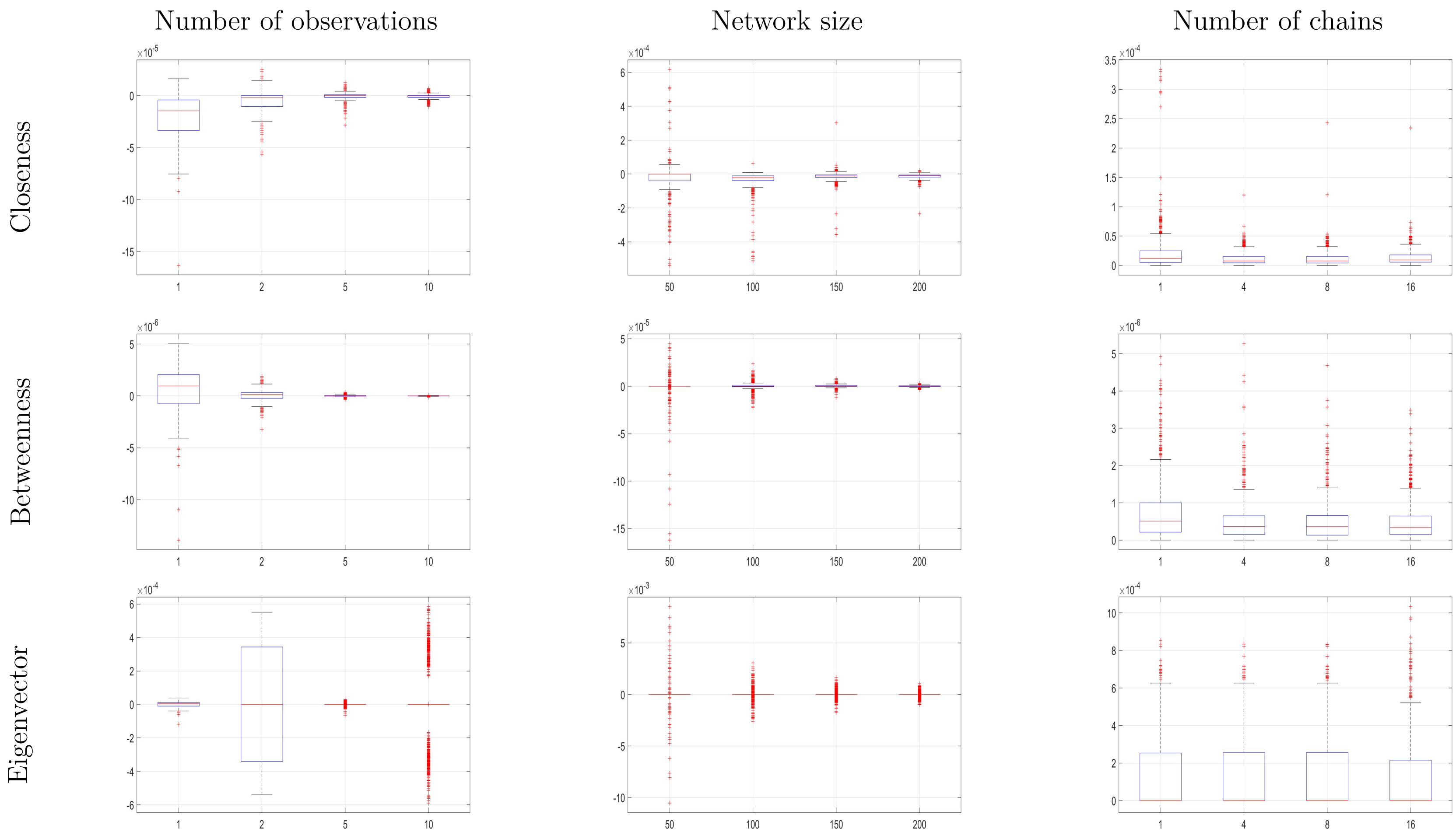
NOTES. Mean values of centrality measures and the p-values of T tests for equality in means are reported. For non-binary characteristics, the means are computed for observations above and below the mean of the distribution. Characteristics of politicians at least one significant difference of network centrality indicators before and after the policy are reported. See Newman [2010] for the definition of the network centrality measures. A precise definition of the variables can be found in Table A.2.

Figure 5: ESTIMATION BIAS - PARAMETERS -



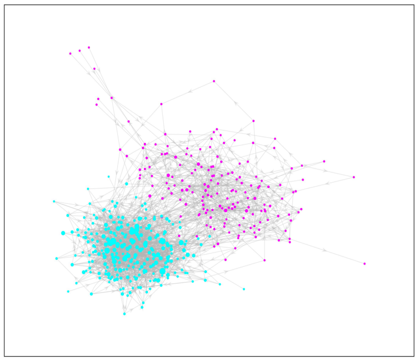
NOTES. X-axis: number of repeated observations of vector E (left panel); network size (center panel); number of MCMCs used for the estimation (right panel). Y-axis: distribution of the differences between the true values of the parameters and the estimated values. True and estimated values are constructed 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 its 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 6: ESTIMATION BIAS - NODE-LEVEL STATISTICS -

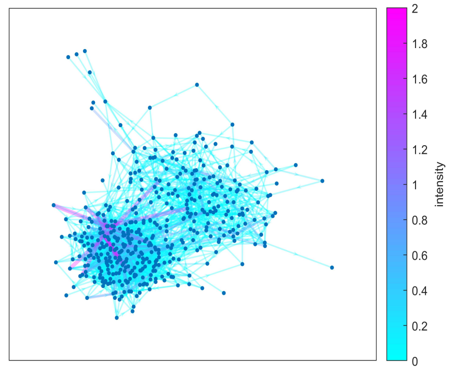


NOTES. X-axis: number of repeated observations of vector E (left panel); network size (center panel); number of MCMCs used for the estimation (right panel). Y-axis: distribution of the differences between estimated and true values of the centralities as defined in Newman [2010]. True and estimated values 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 its 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.

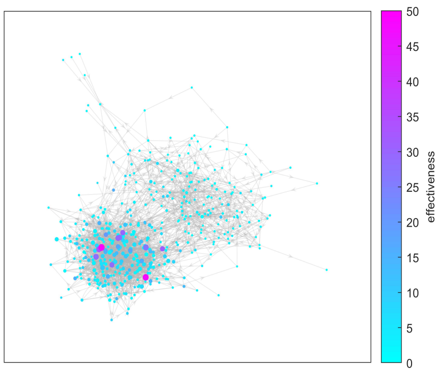
Figure 7: ESTIMATED NETWORK



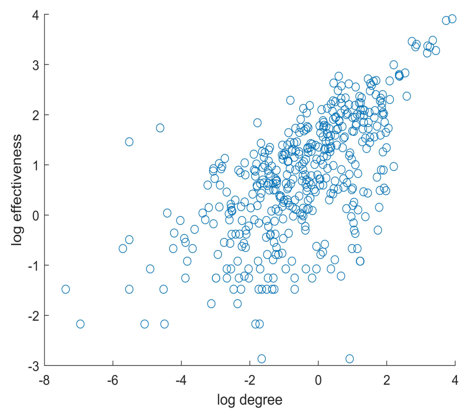
(a) Party



(b) Strength of connections



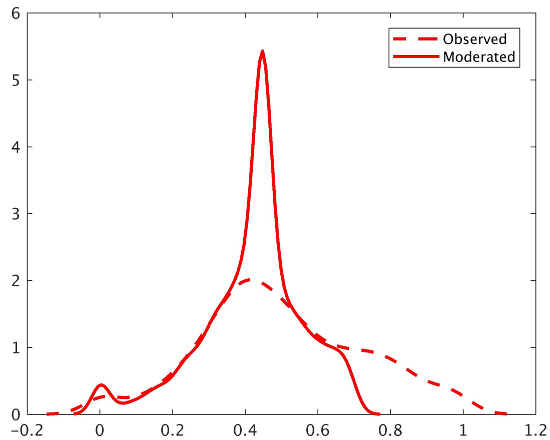
(c) Effectiveness



(d) Degree vs effectiveness

NOTES. The estimated network is derived using the parameter estimates from the 10,000th iteration of the MCMCs for the 111th Congress. In panel (a), the color represents the party of the politician. Red nodes are Republicans and blue are Democrat. In panel (b), the color of links is proportional to g_{ij}, the strength of the connection between politicians i and j. In panel (c), the color of each node is proportional to effectiveness, darker nodes represent more effective politicians. In panel (d), each circle is a politician. The x-axis represents the (log) degree, the y-axis represents the (log) effectiveness. The estimated network is represented with a force-directed layout with five iterations. It uses attractive forces between adjacent nodes and repulsive forces between distant nodes. To ease the visualization, the size of the nodes is equal to the (log) of their degree plus 2. See Fruchterman and Reingold [1991] for more details.

Figure 8: DW IDEOLOGY
- OBSERVED AND COUNTERFACTUAL DISTRIBUTIONS -



NOTES. X-axis: DW ideology . Y-axis: Kernel density estimates. The dashed line represents the observed distribution. The bold line represents the counterfactual distribution where values of DW ideology greater than 0.7 are set equal to the average of the values below 0.7.

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:

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

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

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

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

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

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:

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

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 $i \in M_{M^{-1}(i)}$)

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

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

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

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 $G(\mathbf{g}_{\mathcal{M}\backslash m})$ be the network whit the leaders up to the $(m-1)th$ be $\mathbf{g}_{\mathcal{M}\backslash m} = \{\mathbf{g}_l\}_{l=1}^{m-1}$. 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):

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

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 $\delta < (1/\overline{\theta}) \cdot \left[1/\left(\left(1+\lambda\right)\alpha^{\lambda}\overline{m}\right)\right]^{1/(1+\lambda)}$. 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.

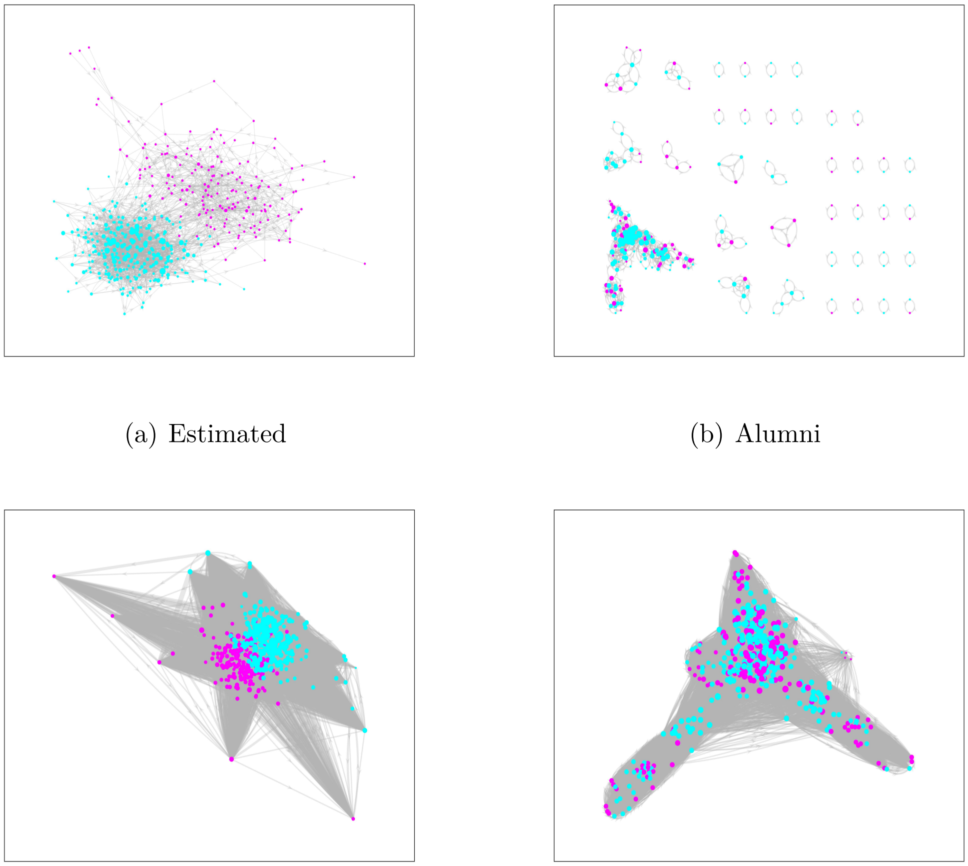
The first one is the number of years served as a member of Congress (*seniority*). As legislators spend more time in Congress, they are expected to become better and more effective at lawmaking. Consistent with the acquisition of skills over time, the second factor is previous *legislative experience*. 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

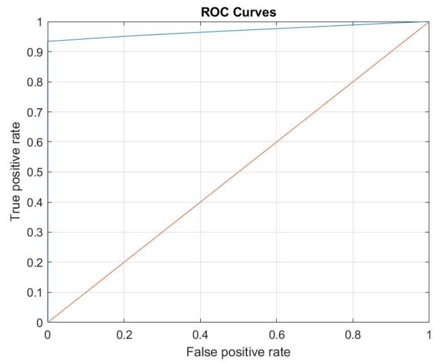
(b) Alumni

(c) Cosponsorship

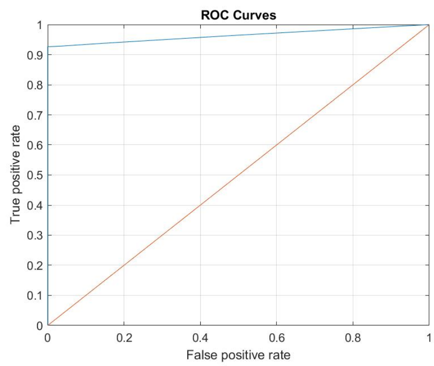
(d) Committee

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



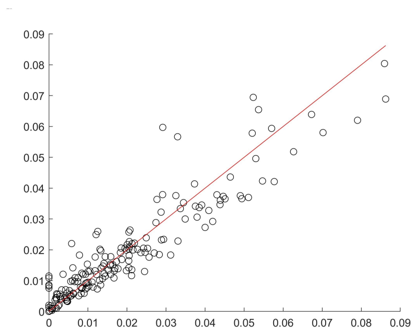
(a) Alumni network



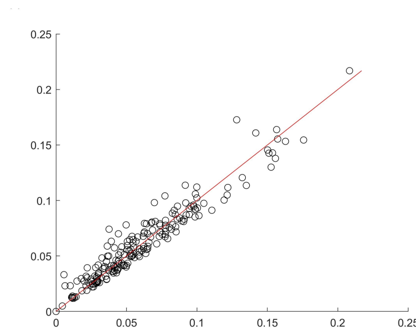
(b) Erdos-Renyi network

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.

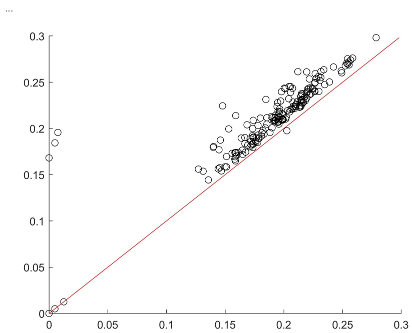
Figure A.3: NODE-LEVEL STATISTICS
- ESTIMATED VS TRUE NETWORK -
ERDOS-RENYI NETWORKS -



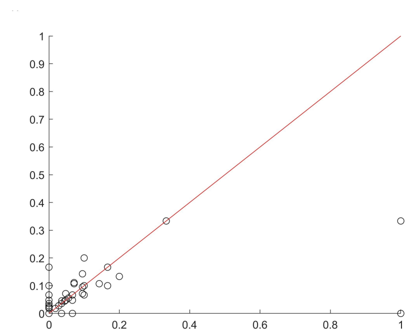
(a) Betweenness



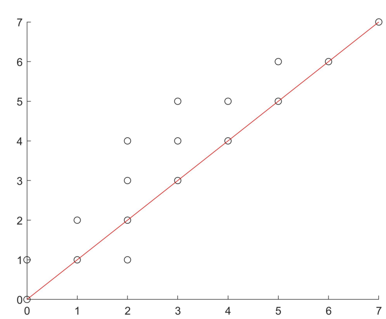
(b) Eigenvalue



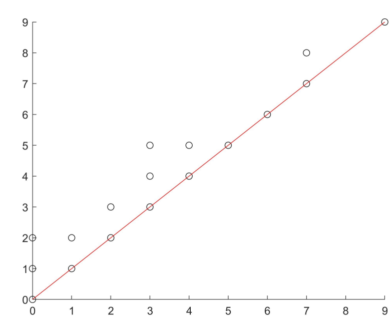
(c) Closeness



(d) Clustering



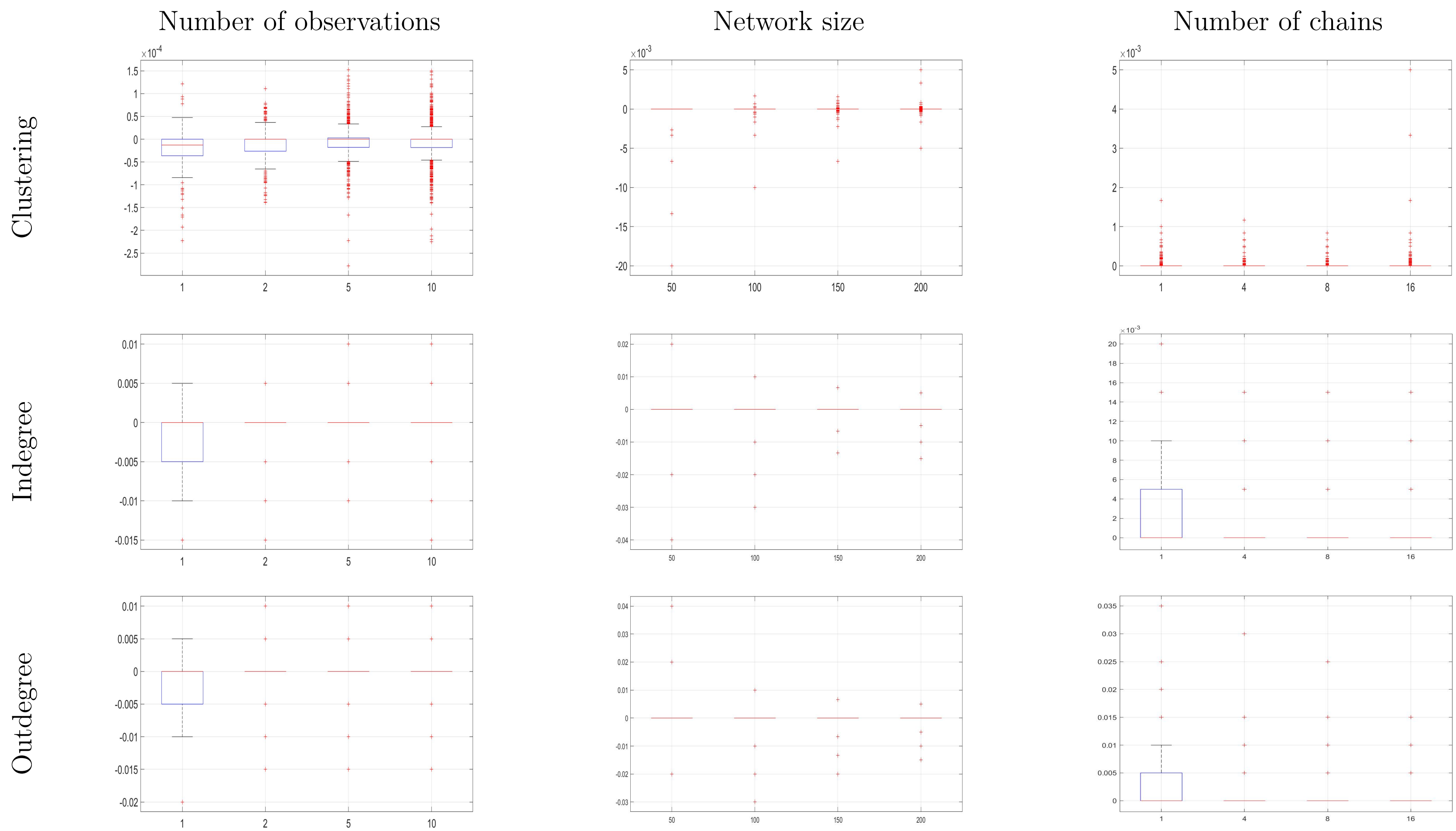
(e) Indegree



(f) Outdegree

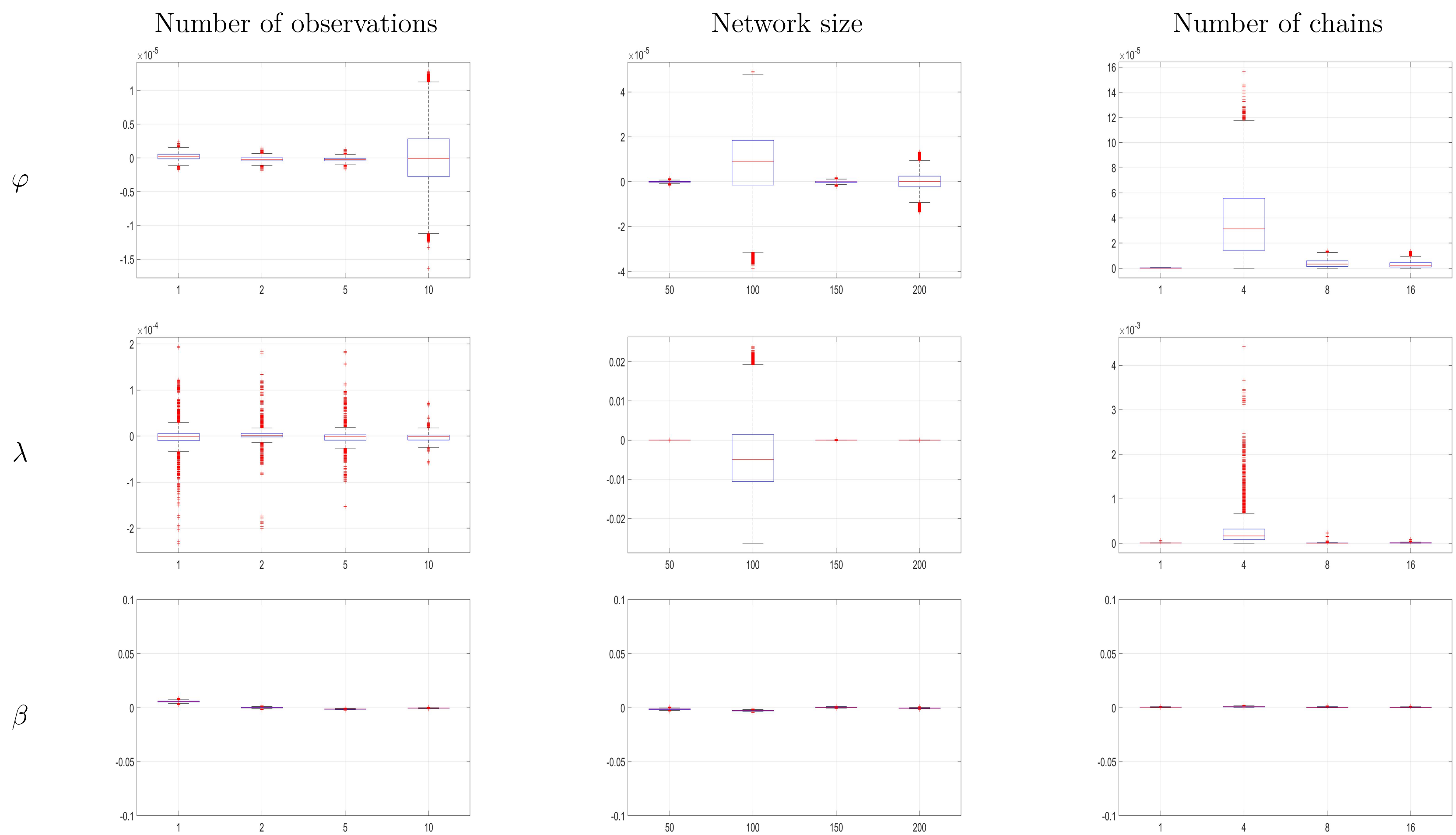
NOTES. X-axis: estimated value of node-level centralities as defined in Newman [2010]. Y-axis: true value of node-level statistic. The true values are the centralities of the true network in which the cost of forming a link depends on the Erdos-Renyi network. The estimated values are the centralities 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 -



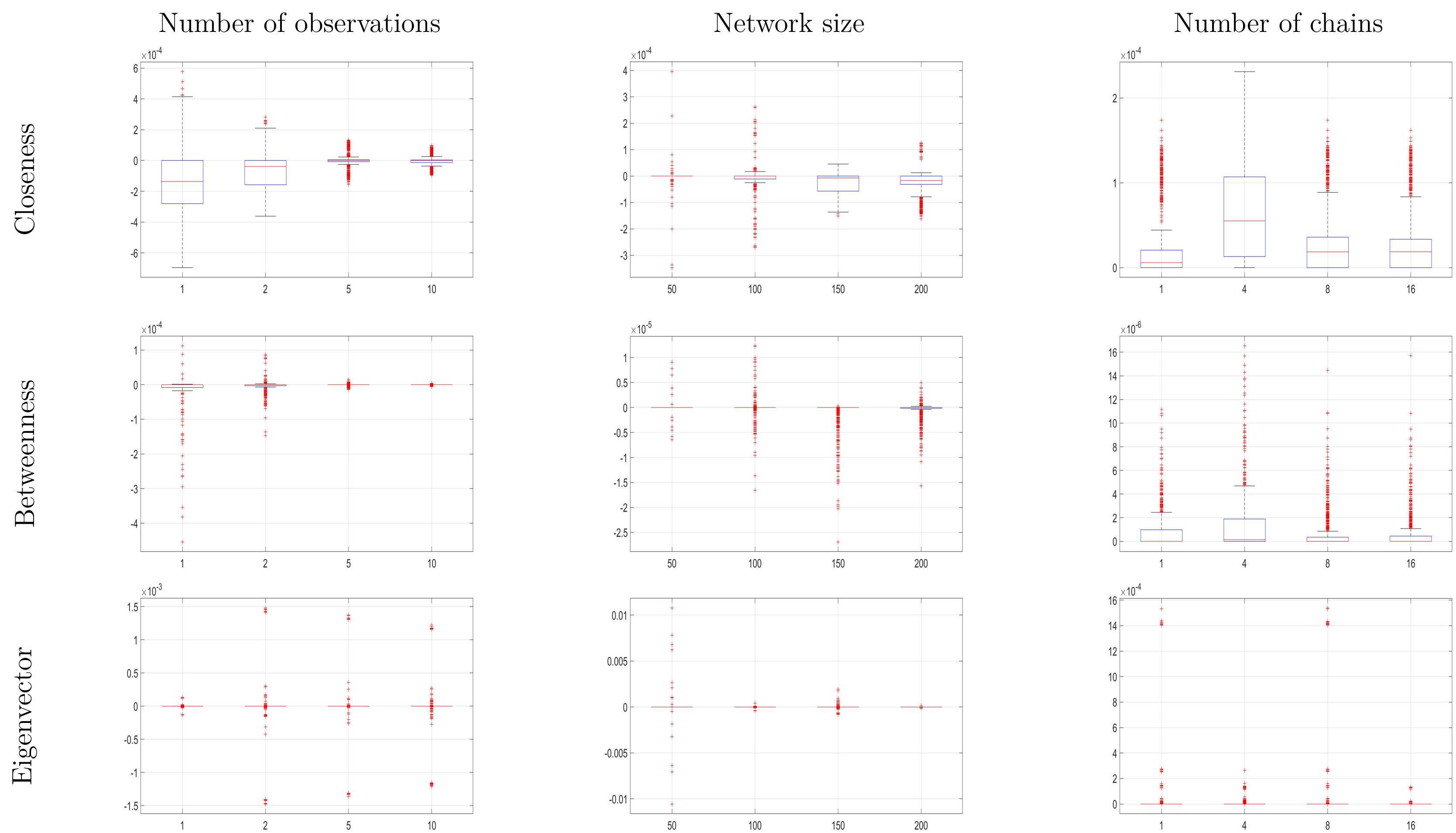
NOTES. X-axis: number of repeated observations of vector E (left panel); network size (center panel); number of MCMCs used for the estimation (right panel). Y-axis: distribution of the differences between estimated and true values of the centralities as defined in Newman [2010]. True and estimated values 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 -



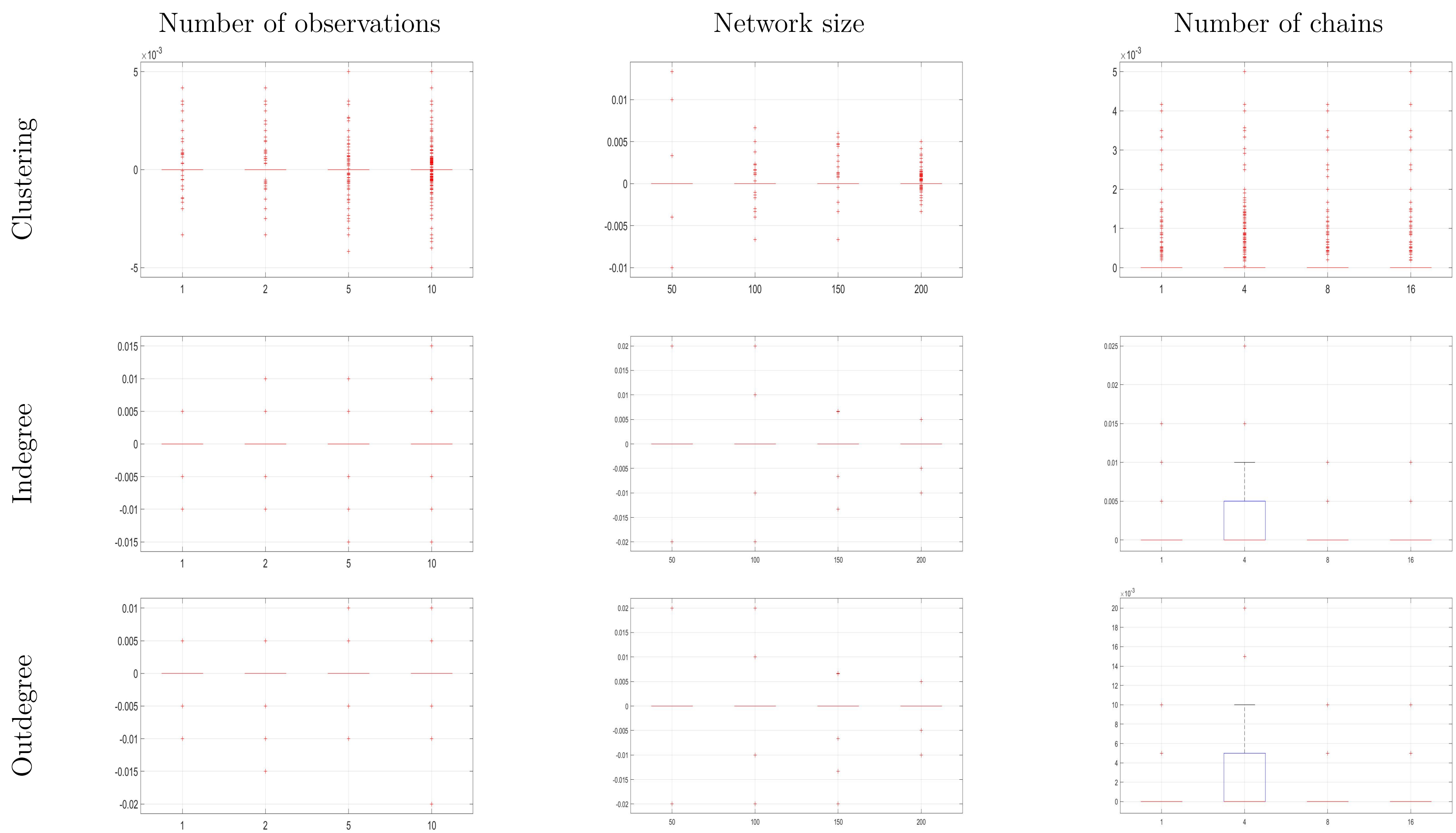
NOTES. X-axis: number of repeated observations of vector E (left panel); network size (center panel); number of MCMCs used for the estimation (right panel); Y-axis: distribution of the differences between the true values of the parameters and the estimated values. True and estimated values are constructed 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 -



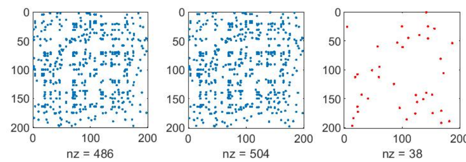
NOTES. X-axis: number of repeated observations of vector E (left panel); network size (center panel); number of MCMCs used for the estimation (right panel). Y-axis: distribution of the differences between estimated and true values of the centralities as defined in Newman [2010]. True and estimated values 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 -

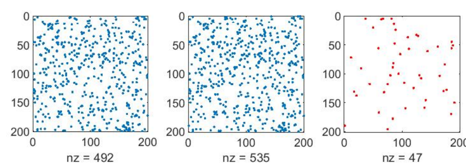


NOTES. X-axis: number of repeated observations of vector E (left panel); network size (center panel); number of MCMCs used for the estimation (right panel). Y-axis: distribution of the differences between estimated and true values of the centralities as defined in Newman [2010]. True and estimated values 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.

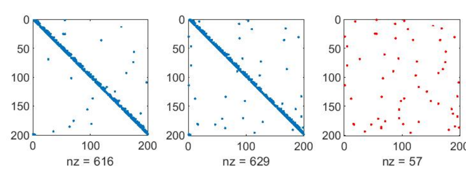
Figure A.8: ESTIMATED VS TRUE NETWORKS - DIFFERENT TOPOLOGIES -



(a) Alumni



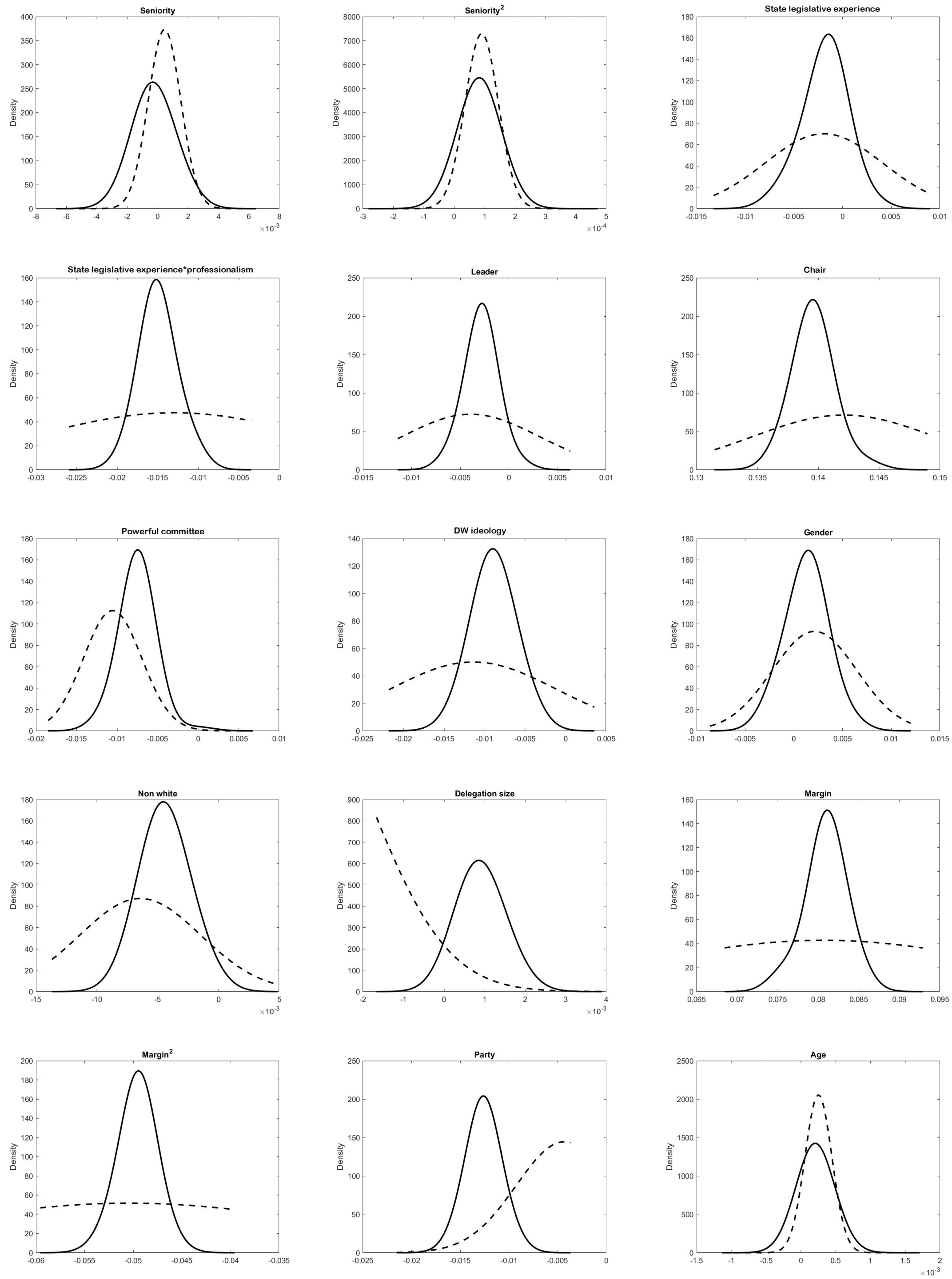
(b) Erdos-Renyi



(c) Circular

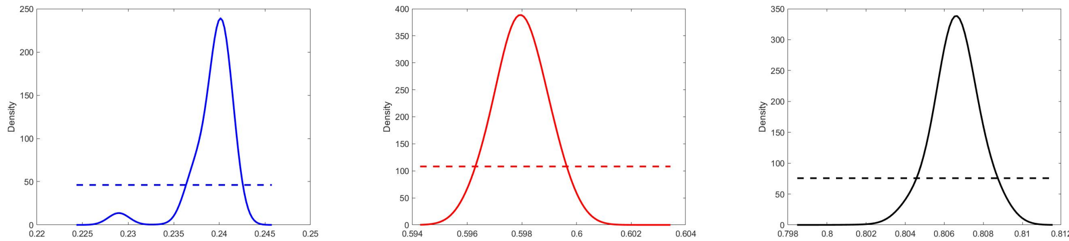
NOTES. Adjacency matrices of the true network, the estimated network (with blue dots) and their difference (with red dots) with n=200. The true and estimated networks are generated as described in Section 8.5. The DGP is described in detail in Section 8.5. The true networks in panels (a), (b) and (c)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

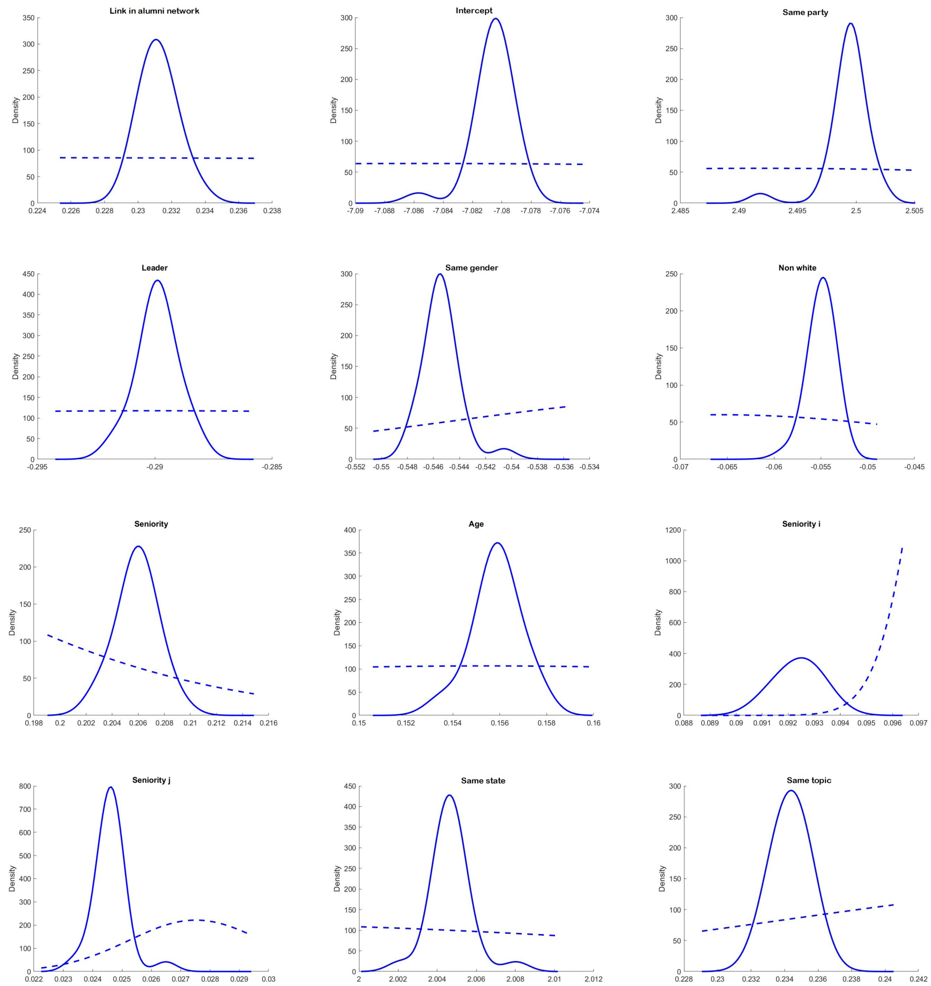
Figure A.10: ESTIMATED POSTERIOR DISTRIBUTIONS
- TARGET VARIABLES -



(a) ρ (b) λ (c) α

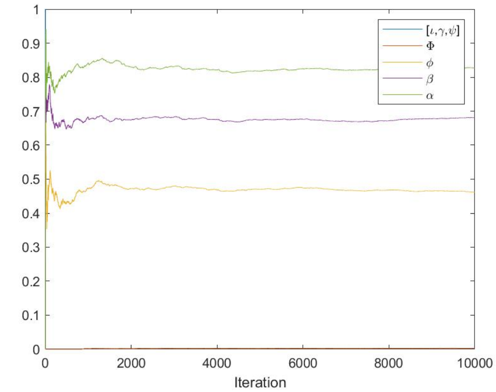
NOTES. X-axis: parameter value, Y-axis: kernel density. The solid line represents the posterior distribution of the parameter estimated by the ABC algorithm, the dashed line depicts the prior distribution.

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



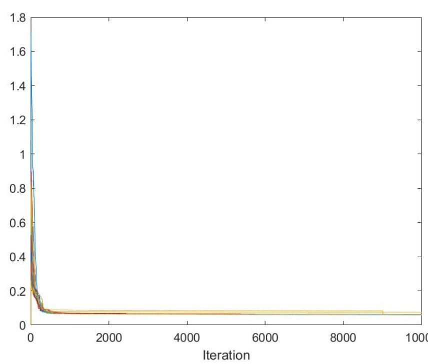
NOTES. X-axis: parameter value, Y-axis: kernel density. The solid line represents the posterior distribution of the parameter estimated by the ABC algorithm, the dashed line depicts the prior distribution.

Figure A.12: ACCEPTANCE RATE AT EACH ITERATION

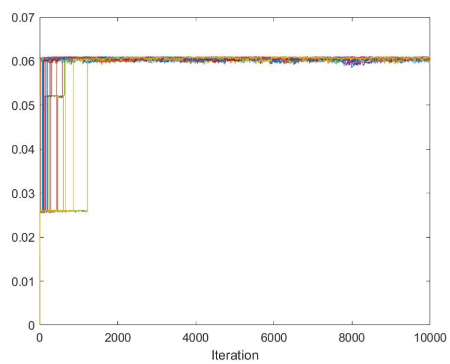


NOTES. X-axis: MCMCs iteration in the second step of the ABC algorithm, Y-axis: acceptance rate. Acceptance rates of each parameter are averaged across the Markov chains.

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



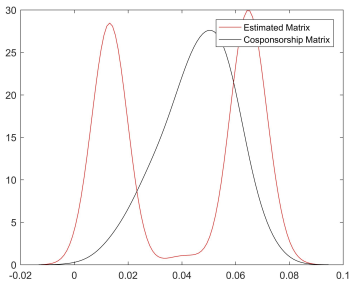
(a) First step



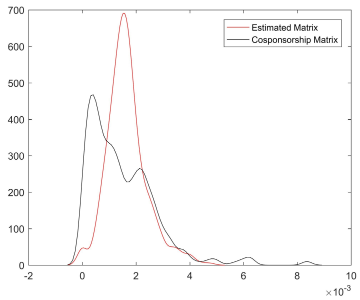
(b) Second step

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.

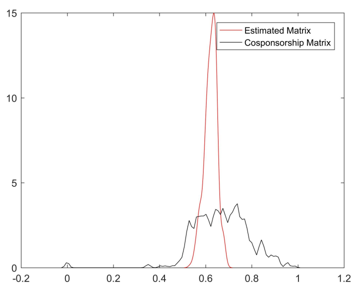
Figure A.14: ESTIMATED VS COSPONSORSHIP NETWORK
- DENSITIES OF NODE-LEVEL STATISTICS -



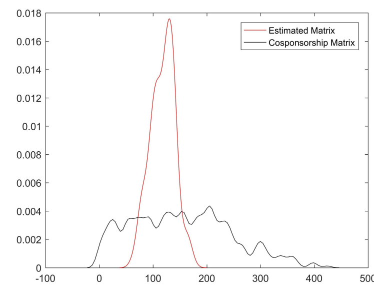
(a) Eigenvector



(b) Closeness

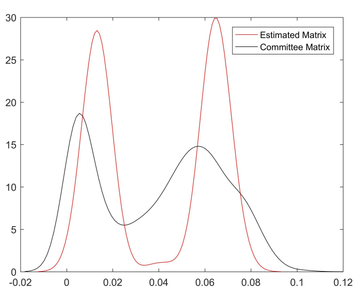


(c) Clustering

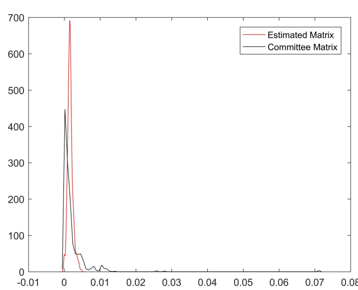


(d) Degree

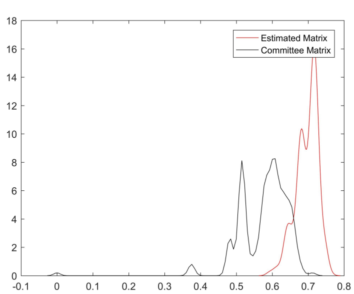
Figure A.15: ESTIMATED VS COMMITTEE NETWORKS
- DENSITIES OF NODE-LEVEL STATISTICS -



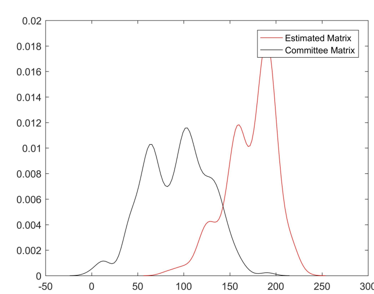
(a) Eigenvector



(b) Closeness

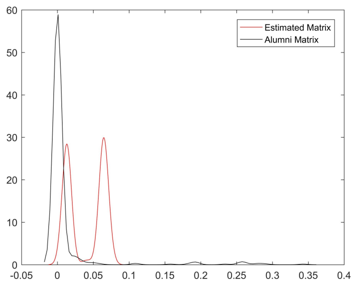


(c) Clustering

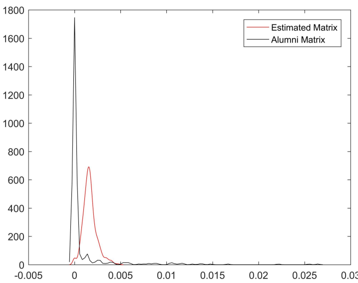


(d) Degree

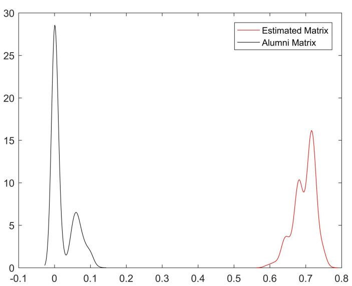
Figure A.16: ESTIMATED VS ALUMNI NETWORKS
- DENSITIES OF NODE-LEVEL STATISTICS -



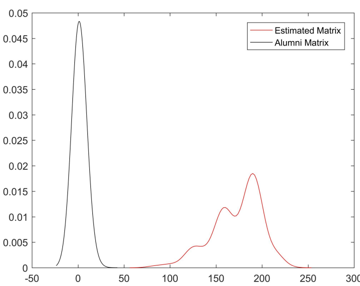
(a) Eigenvector



(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 * State legislative professionalism	-0.0151 ***	-0.0128 ***	-0.0001
	[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.