DISCUSSION PAPER SERIES

DP13845

# ENDOGENOUS SOCIAL CONNECTIONS IN LEGISLATURES

Eleonora Patacchini, Marco Battaglini and Edoardo Rainone

PUBLIC ECONOMICS

*ISSN 0265-8003*

# ENDOGENOUS SOCIAL CONNECTIONS IN LEGISLATURES

*Eleonora Patacchini, Marco Battaglini and Edoardo Rainone*

Discussion Paper DP13845
Published 08 July 2019
Submitted 01 July 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

# ENDOGENOUS SOCIAL CONNECTIONS IN LEGISLATURES


## Abstract

We present a model of the U.S. Congress in which social connections among Congress members are endogenous and matter for their legislative activity. We propose a novel equilibrium concept for the network formation game 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 the model can be structurally estimated by an appropriately designed Approximate Bayesian Computation method. Estimating the model using data from the 109th to 113th U.S. Congresses, we show that social connections are important for legislators' productivities and we identify some of the key determinants of social centralities in Congress.



JEL Classification: D71, D72

Keywords: N/A



Eleonora Patacchini - ep454@cornell.edu
*Cornell University and CEPR*

Marco Battaglini - battaglini@cornell.edu
*Cornell University and CEPR*

Edoardo Rainone - edoardo.rainone@bancaditalia.it
*Bank of Itaky*



Acknowledgements
For useful comments and discussions we thank Coralio Ballester, Vincent Boucher, Antonio Cabrales, Chih-Sheng Hsieh, Angelo Mele, Pedro CL Souza, Mathieu Taschereau-Domouchel and 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), the 2018 Barcelona GSE Summer Forum, and the 2019 Cambridge-INET Conference on Networks. We also thank Neelanjan Datta for excellent research assistance. The views expressed here do not necessarily reflect those of the Bank of Italy.

# 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

This version, May 2019

**Abstract**

We present a model of the U.S. Congress in which social connections among Congress members are endogenous and matter for their legislative activity. We propose a novel equilibrium concept for the network formation game 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 the model can be structurally estimated by an appropriately designed Approximate Bayesian Computation method. Estimating the model using data from the 109th to 113th U.S. Congresses, we show that social connections are important for legislators' productivities and we identify some of the key determinants of social centralities in Congress.

*For useful comments and discussions we thank Coralio Ballester, Vincent Boucher, Antonio Cabrales, Chih-Sheng Hsieh, Angelo Mele, Pedro CL Souza, Mathieu Taschereau-Domouchel and 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 for excellent research assistance. The views expressed here do not necessarily reflect those of the Bank of Italy.

# 1 Introduction

Legislative processes are typically modelled in the political economy literature as bargaining protocols. The prototypical approach to describe legislative bargaining is the Baron and Ferejohn's [1989] model, where a randomly selected proposer makes sequential take-it or leave-it offers to the other legislators until a policy that wins a majority is chosen. In this and other related models, votes are traded in "markets" where only the monetary size of concessions matters.[1] There is no room for politicians' characteristics, their social connections, or even for political parties. There is, however, an important body of work in political science that has stressed the importance of taking into account interpersonal relations and social connections when studying how legislatures work. The history of the U.S. Congress is indeed rich in examples where voting coalitions are shaped by social connections formed inside and outside of the legislative chambers.[2] While the suggestion that social connections should be important in the legislative process is intuitive, little formal work has been done to study them. Can we formally measure their effect on the legislative process? Can we assess what determines whether a legislator is well connected and what this implies for his/her legislative effectiveness?

In answering these questions, the main challenges are that social networks in legislatures (as is the case in most other social groups) are endogenous and not directly observed by outsiders. The standard approach in the empirical literature on network effects is not designed for these environments. The typical approach is to assume that a network is an *observable* realization from an unknown data generation process; and that it can be used to estimate key parameters of the process (for example the cost of link formation and other parameters that matter for the network). To address the questions listed above we need to allow for environments in which networks are *unobservable*: in which, more precisely, only the activities of their members are observable and usable for estimation. This is challenging because models of network formation, even when they ignore the implications for their members' activities, are characterized by significant equilibrium multiplicity and computational complexities.

In this paper, we propose a new theory of legislative behavior in which both the choice of forming social links among legislators and their legislative effort are endogenous. The theory has two stages. In the first stage, legislators target their efforts to form social links with specific other legislators. The effort required to form social links is costly, and bilateral links between two legislators may depend on both of the legislators' efforts, according to a given production function.

[1] We use the term "market" as an analog of a bargaining protocol that is nothing less than a description of how the buyer and seller (of a vote) interact. Other models of legislative bargaining will be discussed in the literature review below. All of these models establish specific market/bargaining protocols in which votes are traded.

[2] We will discuss the political science literature more extensively later in this section.

In the second stage, legislators' effectiveness depends on their legislative efforts and the effectiveness of the legislators with whom links have been established in the previous stage. The preferences of the legislators in the linking process may not only depend on the legislators' observable characteristics, but also on unobservable factors that may be correlated with variables that directly affect the legislators' effectiveness. We use the theory to structurally estimate social links in the 109th to 113th U.S. Congresses using data on legislators' characteristics and activities.[3]

We are able to overcome the complications in solving and estimating the model described above because of two methodological contributions that may have wider applicability in estimating social network effects. First, we introduce a new equilibrium concept that we call *Network Competitive Equilibrium*.[4] 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 make the analysis hard: 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: where 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 fact 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, legislators choose their socialization efforts taking the other legislators' equilibrium effectivenesses as given; these equilibrium effectiveness levels, however, need to be consistent with individual choices. Given this, we show that a Network Competitive Equilibrium can be characterized by a system of nonlinear equations. Moreover, under empirically testable conditions, the equilibrium is unique.

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 (henceforth, ABC), a computational approach that has proven

[3]Measuring legislators' activities is naturally a complicated task. Thankfully, considerable work has been done in the political science literature on this front. For our analysis we rely on the index of *Legislative Effectiveness* constructed by Volden and Wiseman [2014].

[4]Other approaches used in the literature to model complex endogenous network formation process that can be used in empirical analysis include modelling networks as the steady states of stochastic best response dynamics, in which myopic agents select links sequentially; or to adopt cooperative solutions as pairwise stability. We will discuss these approaches and how they relate to our approach in greater detail below and in Section 3.

useful in population genetics and other applications that require large scale models.[5]

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 200 members, even if the networks are dense and we only observe outcomes (here, legislators' effectiveness) in a few periods.[6]

We then use 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 1% 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 0.74% increase in individual effectiveness. Consistent with the endogeneity of the network, moreover, we also find that the elasticity of link formation with respect to other legislators' effectiveness levels is significantly positive.

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 that interpersonal relations bonded long before Congress's election victory are also a key factor, about one-fourth as large as party affiliation. This is in line with the 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]). More importantly, however, 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. Instead, we use a variety of sources of information and let the data tell us what matters. For example, as we discuss in greater detail in Section 5.2, 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.

The underlying assumption in our approach is that agents in our network are marginal and can therefore take the centralities of the other players as given. This approach seems 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. An advantage

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

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

of our approach, however, is that it allows us to obtain sharp and testable predictions for large and complicated networks.

**Related literature** As we discussed above, legislative activity typically is modelled as a bargaining protocol in the political economy literature. The classic reference is Baron and Ferejohn [1989], while alternative bargaining protocols have been presented by Austen-Smith and Banks [1988], Morelli [1999], Baron and Diermeier [2001], Seidmann et al. [2007], Battaglini [2019], and others. In these works, there is little space for the legislators' social connections. In extending the analysis to incorporate social connections, our work relates to two main other strands of literature, one in political science and one in economics.

First, our work relates to a recent literature studying the impact of social networks in Congress on legislative behavior. 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.[7] Peoples [2008], Masket [2008], Rogowsky and Sinclair [2012], Cohen and Malloy [2014] and Harmon et al. [2017] have studied whether social links affect voting behavior. Battaglini and Patacchini [2018] have studied the impact of the legislators' social connections on PAC's contributions. Closer to our work, Fowler [2006], Kirkland [2011], and Battaglini et al. [2019] have studied the relationship between legislators' effectiveness in the U.S. Congress and their social connections. None of these papers, however, model the process of social linking; they all rely upon specific observable social networks as proxies for the true social connections among the legislators.[8]

The second literature to which our work is connected is the economic literature on the estimation of social networks. The traditional approach in the empirical works on endogenous network formation is to interpret *observable* social networks as realizations from unknown data generating processes, and use them for estimating the processes. The issue of estimating an *unobservable* network using other observable economic outcomes has been addressed in the networks literature only by assuming that social connections are exogenous. The complication here is that the number of network connections is much higher than the number of economic outcomes that can be used for the estimation (in our specific application, the legislators' performance in Congress): with $n$ legislators observed for $T$ sessions, the first is on the order of $n^2$, the second just of order $n \cdot T$. De

[7] Earlier work includes Rice [1927, 1928], Routt [1938], Eulau [1962]. See Victor et al. [2016] and Battaglini and Patacchini [2019] for recent surveys.

[8] The observable social connections that have been used include, among others, the proximity network, in which two legislators are linked if they sit next to each other in a legislative session or if they have nearby offices; the cosponsorship network in which two legislators are linked if they have cosponsored each other's bills; and the alumni network, in which two legislators are linked if they have attended the same educational institution. See Battaglini and Patacchini [2019] and Victor et al. [2016] for surveys of this literature.

Paula et al. [2018] use high-dimensional estimation techniques to estimate social networks, which can bypass the dimensionality problem.[9] While this and other related approaches are versatile and can be applied to many environments with minor changes,[10] they have two main limitations. First, they typically require assuming an exogenous linear model of behavior, thus not allowing for the endogeneity of the network. Second, and more importantly, they require assuming that a social network is sparse and 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 network are generally dense networks that cannot be observed repeatedly because elections are held every few years.[11] Our approach, instead, is to rely on the structure generated by our endogenous model of link formation to reduce the dimensionality of the estimation; and then to use the observable outcome from just one or 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 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.[12] 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 only leads only to partial identification under strong sparsity conditions on the network.[13] 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 details.

Closer to our approach in modelling network formation is the work by Acemoglu and Azar [2018],

[9] 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. The properties of the estimators with missing data have also been studied by Shalizi and Rinaldo [2013], Handcock and Gile [2010], Koskinen et al [2010], and others. See De Paula [2017] for a comprehensive discussion.

[10] For related approaches, see also Manresa [2016] who, however, only considers exogenous peer effects; and Battaglini et al. [2018], who follow a graphical lasso approach to identify the network.

[11] Elections for the U.S. Congress are held every two years. While reelection rates are very high (about 90-95% of Congress 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% of first-term Congress members do not seek reelection (Glassman and Wilhelm [2017]). It is therefore problematic to assume social networks in Congress persist for more than a few Congresses.

[12] 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 [2018] have characterized the distribution of networks emerging from alternative sequential models.

[13] 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. We refer to De Paula [2017] for a comprehensive survey of this literature.

who study the endogenous choices of firm input combinations in a competitive environment.[14] In their model marginal productivities of inputs depend on the entire input/output network, but are taken as given by firms when chosing suppliers. Contrary to us, they 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.

Two other papers in the networks literature deserve special mention. The first is Boucher [2018], who also uses an Approximate Bayesian Computation approach similar to ours to estimate a model of endogenous network formation. In contrast to our approach, he aims at estimating the probability distribution over networks under the assumption that a network realization is observed. The sequential model of social link formation underlying this work does not have a tractable analytical characterization of the equilibria, thus it does not give sufficient statistics for the approximation of the likelihood function. Boucher [2018] suggests a set of summary statistics that allows him to estimate homophily in a network of high school friendship based on Add Health data. The second related paper is Canen et al. [2017], who adapts a model by Cabrales et al. [2011] to estimate unobservable social efforts in the U.S. Congress. In contrast to this 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").[15] The analysis, moreover, is based only on the total number of cosponsored bills by a legislator, that is used as an empirical proxy for social effort; and on a one-dimensional index of roll call data and floor speeches, that is used as an empirical proxy for legislative effort.

The remainder of this paper is organized as follows. Section 2 presents our model of legislative behavior formation of social connections. Section 3 defines our equilibrium concept, the Network Competitive Equilibrium, and characterizes it. Section 4 defines the econometric specification of the model, discusses the estimation method by Approximate Bayesian Computation Method, and presents a set of simulations to explore the performance of our approach in finite samples and as we change the environment of the estimation. Section 5 estimates the model using data from the 109th to 113th U.S. Congresses. Section 6 presents a few extensions to the model and discusses the robustness of the results. Section 7 concludes.

[14] Other recent but less related models of endogenous production networks are presented by Oberfield [2018] and Taschereau-Dumouchel [2019].

[15] 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

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 cares to implement. The goal of each legislator is to maximize their legislative effectiveness, measured by the probability of implementing the project.[16] 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:

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

The Cobb-Douglas 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 the higher is the effectiveness of the legislators socially connected to $i$, the higher is $i$'s effectiveness; 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, $\varepsilon_i$, is a factor idiosyncratic to $i$ that contributes to $i$'s efficacy independently from his/her connections or effort. We assume it is observed by the players, 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**. $\varphi \cdot \overline{g}^{\alpha} \cdot \overline{l}^{1-\alpha} + \overline{\varepsilon} < 1$.

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 $G$

[16] 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], Weingast [1979], Shepsle and Weingast [1981], Weingast and Marshall [1988] among others).

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 $j \in N \backslash i$ he or she wishes to establish a link $g_{i,j}$. A link $g_{i,j}$ is established only if $j$ approves it, but it only depends on $i$'s effort. If a social link from $i$ is approved by $j$, the cost of establishing it with intensity $g_{i,j}$ is given by:

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

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 to coordinate actions, or the cost of the 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 district, whether they have the same party affiliation, sex, social, or educational background (for example if they attended the same educational institutions, etc.). 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 2.3, where we develop the empirical analysis of the model. While $\Theta$ and $\mathcal{M}_i$ are exogenous in the theoretical model, in the empirical analysis we will let the data identify them.

In the socialization process described above, the ability of $i$ to establish a link with $j$ depends only on $i$'s effort and on $i$ 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 $i \in \mathcal{N}$.[17]

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

Note that a simple condition that guarantees both Assumption 1 and Assumption 2 are satisfied

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

is that the parameter controlling the social spillovers $\varphi$ is sufficiently small.

The type $\omega_i$ 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 $\Omega$ 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 $l : \Omega \times G \to [0, \overline{l}]$, mapping the social network and $i$'s type to an effort level.

# 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 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, in addition to a straightforward direct effect, 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 $E_j$s of $j$s who are connected to $i$; these changes may affects $E_j$ if $i$ 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 we have when studying a general equilibrium in an exchange economy in 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 legislators' levels of effectiveness. We can

therefore introduce a *Network Competitive Equilibrium* (henceforth, NCE) as follows:

**Definition 1**. *Legislators' effort levels* $l = \{l_1, ..., l_n\}$*, legislative effectiveness* $\mathbf{E} = \{E_1, ..., E_n\}$ *and the social matrix* $G = (g_{i,j})_{i,j \in \mathcal{N}}$ *constitute 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' effectivenesses corresponds 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. While the concept is introduced in the context of this model, it seems the idea of a NCE should have a more general applicability to network formation in many other environments. We will explore some properties of this equilibrium definition below.

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. A key aspects 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 16, plotting our estimated network and other observed networks often used to study social networks in Congress (such as the cosponsorship and committee networks).[18] 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,

[18]The cosponsorship network has been used to study social connections in the U.S. Congress by Fowler [2006], Kirkland [2011], Parigi and Sartori [2014] among others. The committee network has been used by Porter et al. [2005].

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.[19] 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.[20]

While the NCE can conceptually be applied to any networks, its applicability is probably not appropriate for simple environments with few links, where it is probably 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 it 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.

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

## 3.2 The choice of effort at $t = 2$

Substituting in the solution to the maximization of (4) into (1), we obtain that the equilibrium levels of legislative effectiveness for a type $i \in N$ are given by:

$$E_i = \varphi \left( \frac{(1-\alpha)\varphi}{c} \right)^{\frac{1-\alpha}{\alpha}} \cdot \sum_{j=1}^{n} g_{i,j} E_j + \varepsilon_i. \tag{4}$$

These equations can be expressed in matrix form as:

$$[I - \delta \cdot G] \cdot \mathbf{E} = \varepsilon \tag{5}$$

where $\delta = \varphi \left( (1-\alpha)\varphi / c \right)^{\frac{1-\alpha}{\alpha}}$ and $\varepsilon$ 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 (5), a legislator's effectiveness coincides

[19] For empirical analysis based on this approach, see Christakis et al. [2010], Konig [2016], Mele [2017], Badev [2017] and Boucher [2018], among others.

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

with his weighted Bonacich centrality in $G$, with weights given by the natural "effectiveness of each legislator" $\varepsilon_i$ and the discount factor $\delta$.[21] However, the 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 $i$ is easily determined by substituting the optimal effort levels and $E_i(G,\varepsilon)$ in (4):

$$U^i(G,\varepsilon) = \alpha\delta \sum\nolimits_{j=1}^{n} g_{i,j} E_j(G,\varepsilon) + \varepsilon_i. \tag{6}$$

Legislator $i$ will choose the links $\mathbf{g^i} = (g_{i,1}, ..., g_{i,n})$ that maximize (6) 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.[22] It follows that legislator $i$ chooses his links solving:

$$\max_{\mathbf{g^i}} \left\{ \sum\nolimits_{j=1}^{n} \left[ \alpha\delta \cdot g_{i,j} E_j(G,\varepsilon) - \frac{\phi}{(1+\phi)} \left( \frac{g_{i,j}}{\theta_{i,j}} \right)^{1+\frac{1}{\phi}} \right] \right\}. \tag{7}$$

Combining the solution of (7) with (5), we have:

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

$$E_i^* = \delta \cdot \sum\nolimits_{l \in \mathcal{N}} \left( g_{i,l}^* E_l^* \right) + \varepsilon_i \tag{8}$$

$$\text{and } g_{i,j}^* \leq (\theta_{i,j})^{1+\phi} \left( \alpha\delta E_j^* \right)^{\phi} \ (= \text{ for } g_{i,j}^* \leq \overline{g}) \tag{9}$$

*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

[21] The standard definition of the Bonacich centrality with discount factor $\nu$ is $\mathbf{E} = [I - \nu \cdot G]^{-1} \mathbf{1}$ (See Bonacich [1987]). The weighted Bonacich with weights $\mathbf{A}$ is defined as $\mathbf{E} = [I - \nu \cdot G]^{-1} \mathbf{A}$ (See Ballester et al. [2006]).

[22] 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.

in $n$ equations and $n$ variables:

$$E_i^* = \alpha^{\phi} (\delta)^{1+\phi} \sum\nolimits_{l \in \mathcal{N}} (\theta_{i,l} E_l^*)^{1+\phi} + \varepsilon_i \tag{10}$$

The legislators' effectivenesses are no longer representable by a linear system of equations as in the familiar "Bonacich" representation of (5). 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 $(E_j)^{\phi}$. The true link between $E_i$ and $E_j$, therefore, is no longer linear: $E_i$ will be a function of $(E_j)^{1+\phi}$.

To interpret (8) and (10), 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 $\epsilon_{g_{i,j},E_j} = \phi$.[23] As $\phi \to 0$, 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 $\phi \to 0$, there is a unique equilibrium in which effectiveness* coincides *with the Bonacich centralities: $\mathbf{E} = [I - \delta \cdot \Theta]^{-1} \cdot \varepsilon$, where $\Theta$ is the $n \times n$ matrix with generic term $\theta_{i,j}$.*

When $\phi > 0$, instead, changes in the equilibrium effectiveness imply changes in the links. The higher is $\phi$, the more links polarize around the most effective legislators. To see the implications of a positive $\phi$, let $\phi \to \infty$. 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 a symmetric environment in which legislators are symmetric (so $\varepsilon_i = \varepsilon$ for all $i$) 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 $g_{i,j} = \overline{g}$ if $\theta_{i,j} > 0$. We have:

**Example 2.** *As $\phi \to \infty$, 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 $\varepsilon \in [(1 - 2\delta\overline{g})/(\alpha\delta), 1/(\alpha\delta)]$.*

The network structure described in Example 1, where links are inelastic to the level of effectiveness of the legislators (i.e. $\phi = 0$) 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 $\varepsilon$, 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 (10). This property is implied by the following assumption

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

on the fundamentals:

**Assumption 2**. $\overline{g} > (\alpha\delta)^{\phi}\,\overline{\theta}^{1+\phi}$.

Under Assumptions 1 and 2, it is easy to state a sufficient condition for the existence of a unique equilibrium for $\delta$ sufficiently small. Define $\overline{\theta} = \max_{i,j\in\mathcal{N}} \theta_{i,j}$. We have:

**Proposition 2.** *For* $\delta < \frac{1}{\overline{\theta}}\left[1/\left((1+\phi)\,\alpha^{\phi}\overline{m}\right)\right]^{1/(1+\phi)}$, *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 $\phi$ obviously violates this condition but still admits a unique equilibrium under the conditions specified above.

# 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 it as $X_{i,r} = (X_{i,r,1}, ..., X_{i,r,l})$. 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 $H_r = (h_{i,j,r})_{i,j\in\mathcal{N}}$. In the following, $H_r$ will be the alumni network, in which $h_{i,j,r} = 1$ if $i$ and $j$ have attended the same educational institution in overlapping periods, and $h_{i,j,r} = 0$ otherwise.[24] We will discuss these variables in greater detail in the next section.[25]

Propositions 1 shows that, in equilibrium, effectiveness solves (10). To bring this system of equations to the data, we assume that $\varepsilon_{i,r} = (\mathbf{X}_{i,r})'\beta + \zeta_r + \epsilon_{i,r}$, where $\beta$ is a vector of coefficients, $\zeta_r$ is a Congress fixed effect and $\epsilon_{i,r}$ is a random variable. We therefore have:

$$E_{i,r} = \alpha^{\phi}\,(\delta)^{1+\phi} \sum\nolimits_{l\in r} \left(\theta_{i,l,r} E_{l,r}\right)^{1+\phi} \quad + \mathbf{X}'_{i,r}\beta + \zeta_r + \epsilon_{i,r} \tag{11}$$

[24] 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. [2019]. While we will not assume it to be relevant, we will use it as input of the analysis as described below.

[25] We omit the index for the Congress in $h_{i,j}$ because in the following analysis the exogenous adjacency matrix will not depend on $r$. In general, it might also depend on $r$.

In (11), 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:

$$P\left(\theta_{i,j,r}\,|\lambda_{i,j,r}\right) = \left(\frac{e^{\lambda_{i,j,r}}}{1+e^{\lambda_{i,j,r}}}\right)^{\theta_{i,j,r}} \left(\frac{1}{1+e^{\lambda_{i,j,r}}}\right)^{1-\theta_{i,j,r}} \tag{12}$$

with:

$$\lambda_{i,j,r} = \iota + \gamma h_{i,j,r} + \sum\nolimits_{l} g(X_{i,r,l}, X_{j,r,l})\psi_l \tag{13}$$

In (13) $\iota$ is a constant; $\gamma$ and $\psi_l$ are parameters to be estimated; and $g(\cdot,\cdot)$ is a distance function. The parameter $\lambda_{i,j,r}$ 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 $g(\cdot,\cdot)$. The specification in (13) 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.[26]

In estimating our model, there are three parameters that are of special interest: $\varphi, \phi$, and $\alpha$. The parameter $\varphi$ measures the importance of the social spillovers as is typical of the literature on network effects (see, for example, Calvò-Armengol et al. [2009]). As discussed above, $\phi$ is the elasticity of the link formation with respect to the equilibrium level of effectiveness of the other legislators. This parameter is key to test of whether the endogeneity of the network formation is a significant factor in shaping the legislators' effectiveness. Because the model with an exogenous network (i.e. with $\phi = 0$) is nested in the more general model, we can test if allowing $G$ to be endogenous improves the fit of the model. Finally, $\alpha$ measures the elasticity of $\mathcal{E}_i = E_i - \varepsilon_i$ with respect to social connectedness and effort. The variable $\mathcal{E}_i$ can be interpreted as the endogenous component of $i$'s effectiveness, generated by social connections.

## 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 $\omega$, propose a move to $\omega'$ according to a transition

[26] The functional form assigned to $g(\cdot,\cdot)$ in our empirical application is the absolute value of differences between individual characteristics, which captures homophilic or heterophilic behavior. The logistic function is one of the most popular in dyadic link formation (see for example Graham [2017]). Alternative functional forms and specifications (even not binding $\theta_{i,j,r}$ to be binary) can be used. The estimation method proposed in the next section does not impose particular limitations on these choices.

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 $\omega'$ in the model.

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

Under suitable regularity conditions, the limiting stationary distribution of the chain described above is equal to the conditional 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 Fu and Li [1997] and Weiss and Haeseler [1998] and Plagnol and Tavare [2004] 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 $\omega$, propose a move to $\omega'$ according to a transition kernel $q(\omega \to \omega')$.

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

B3. If $\rho(\mathbf{E}', \mathbf{E}) < \nu$, proceed to the next step otherwise return to the first step. Here, $\mathbf{E}$ is the observed vector, $\rho(\mathbf{E}', \mathbf{E})$ is a norm between $\mathbf{E}'$ and $\mathbf{E}$, and $\nu$ 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 $\omega'$ with probability $h$, else remain at $\omega$; go to the first step.

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

When, as in our case, $\mathbf{E}$ is high dimensional, two problems may arise. First, for practical purposes it may be necessary to replace $\rho(\mathbf{E}', \mathbf{E}) < \nu$ in step B3 with $\rho(\eta(\mathbf{E}'), \eta(\mathbf{E})) < \nu$, where $\eta(\mathbf{E})$ is a vector of summary statistics of **E.** However, this reduces the information that is used because it is hard to identify a set of sufficient statistics (see Robert et al. [2011]). Second, the algorithm may be slow when the evaluation in step B2 is computationally onerous. We can, however**,** bypass both of these problems. First, the characterization of Proposition 1 allows us to drastically reduce the dimensionality of the problem, which is fully characterized by a system of "only" $n$ equations (despite the endogenous network being of dimension $n^2$). This allows us to directly

use the observed vector $\mathbf{E}$, rather than a set of statistics of it. Second, the explicit equilibrium condition in (11) allows us to bypass the need of solving the nonlinear system of $n$ equations by simply evaluating how well the the proposed vector of parameters solves them in correspondence to the observed vector $\mathbf{E}$. Define the vector:

$$z(\mathbf{E},\omega) = \mathbf{E} - \varepsilon - \alpha^{\phi}\delta^{1+\phi} \cdot \Theta \cdot \mathbf{D}(\mathbf{E},\phi)$$

where $\mathbf{E} = (E_1, ..., E_n)'$ is the observed vector of empirical efficacies, $\mathbf{D}(\mathbf{x},\phi) = ((x_1)^{1+\phi}, ..., (x_n)^{1+\phi})'$ and $\Theta$ is a $n \times n$ matrix with $i, j$ element equal to $(\theta_{i,j})^{1+\phi}$. Let $\rho(z(\mathbf{E},\omega))$ be the norm of this vector. Our modified algorithm is:

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

C2. If $\rho(z(\mathbf{E},\omega')) < \nu$, 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 $\omega'$ with probability $h$, else remain at $\omega$; go to the first step.

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* $\mathcal{D}_\nu = \{\omega | \rho(z(\mathbf{E},\omega)) \leq \nu\}$.

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 $\nu$ to zero. The details of how our ABC algorithm is implemented in practice are described in the Appendix (Section 8.6).

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]). The interpretation of this posterior is, moreover, straightforward. Bayes rule allows to use the observations to update the probabilities of the events associated to the sigma algebra of the minimal sufficient parameter space.[27] If the parameter space in the Bayesian model is not minimal, the conditional probabilities

[27] 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 $\sigma$-field 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.

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 parameter for the type of datasets that we use in the empirical application in Section 5. Regarding the empirical analysis of Section 5, Figures A.2 - A.4 in the supplementary online Appendix show that there is significant updating of information in the chains, which result in posterior that are markedly different from the priors.

## 4.3 Monte Carlo Simulations

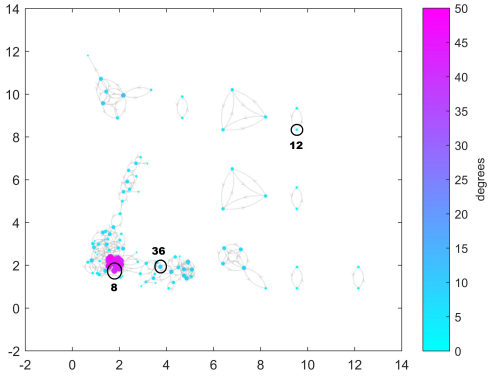
In this section, we use simulations to investigate the performance of our estimation approach. We first propose two examples to illustrate its ability to recover the underlying structural parameters and key features of the network topology. We then systematically study how the performance of our approach changes as we change important features of the setting, such as the sparsity of the networks, their size, the number of periods for which we observe the outcomes, and the elasticity of network formation, among others.

### 4.3.1 Two benchmarks

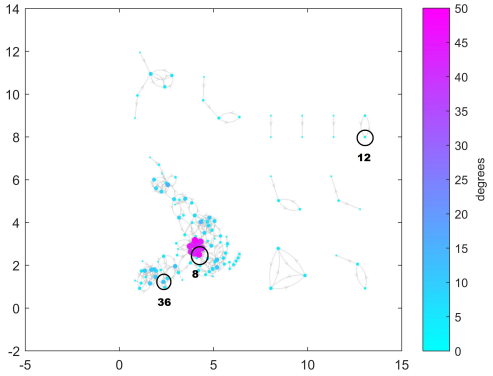
**A real network** There are a number of parameters determining the shape of the social network in our model that need to be chosen in any simulation. An important ingredient is the observed adjacency matrix that enters in the cost function (3), i.e. the observed factors affecting the cost of establishing a link from legislator $i$ to legislator $j$. In our first set of simulations, we use a real network as a basis to simulate $H = \{h_{i,j}\}$. Specifically, we generate $H$ using the alumni network, the same network that we will use in 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 within four years and zero otherwise.[28]

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

Figure 1: REAL NETWORK ESTIMATION
- GOODNESS OF FIT -




(a) Estimated




(b) True

NOTE. Panels (a) and (b) represent the estimated and true network respectively. The true network is generated using equations (26)-(27) and data from the alumni network. The connections of 200 politicians from the 111th Congress extracted at random are considered. The 111th Congress is randomly selected. Results are robust to different draws. The DGP is described in detail in Section 8.5. The estimated network is derived using the parameter estimates at the last iteration of the MCMC. The first of $\bar{c} = 5$ networks is visually represented using the force-directed layout algorithm with five iterations. 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.

We use a given set of parameters $(\iota, \gamma, \psi_l, \alpha, \phi, \varphi)$, randomly generated characteristics $X$, and $H$ as inputs in (12) and (13) to compute $\theta_{i,j}$ and extract the true network $G = \{g_{i,j}\}$ from (9). The

true network $G$, along with realizations of the other random variables $X$, are then used in (8) to generate the observed vector of effectiveness $\mathbf{E}$. We simulate a situation in which five Congresses are observed, which corresponds to simulating five vectors $\mathbf{E}$. The parameters $(\iota, \gamma, \psi_l, \alpha, \phi, \varphi)$ 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 diameter, the average distance of the politicians' cosponsorship network, and the betweenness of the committee network (in which two legislators are linked if they belong to the same committee).[29] Panel (b) of Figure 1 illustrates a simulated network. The estimation procedure consists of using the simulated $\mathbf{E}$, $H$, and $\mathbf{X}$ to estimate $(\iota, \gamma, \psi_l, \alpha, \phi, \varphi)$, and thus the unobserved $G$. Because we know the true $(\iota, \gamma, \psi_l, \alpha, \phi, \varphi)$ in these exercises, we can evaluate the performance of the estimation technique.

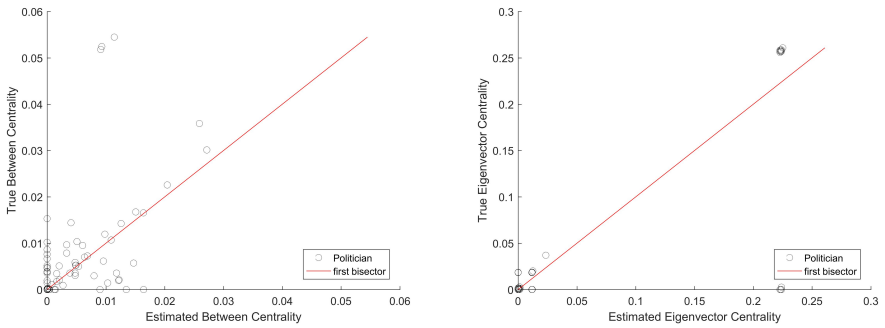
To illustrate the performance of the network estimation, in Figure 1 we show a graphical comparison of the real network (Panel (b)) versus the estimated network (Panel (a)) for one of the five simulations (results for the remaining four simulated networks are similar). The pictures show that the estimated and original networks are, naturally, not identical but remarkably similar in terms of topological structure.[30] 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 highlight three specific nodes, one in the center, one in the periphery, and one in the extreme periphery. Their respective positions in the real 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.1.[31] Again, the picture shows that almost all of the links are correctly predicted.

[29] The details of how we have calibrated the remaining parameters and other details regarding the implementation are presented in Section 8.5 in the Appendix.

[30] In our model, a network depends on deterministic factors and random components affecting the cost of forming a link $\theta_{i,j}$, see Section 4.1 and Appendix 8.5 for the details of the implementation. The network illustrated in Figure 1 is a random realization from the calibrated model using the estimated parameters. The results are, however, representative as similar qualitative results are obtained independently from the draws.

[31] The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. 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 better is the estimation.

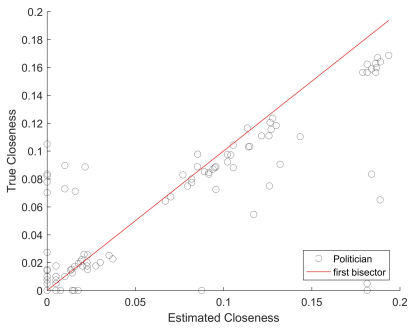
Figure 2: NODE-LEVEL STATISTICS
- ESTIMATED VS TRUE REAL NETWORK -



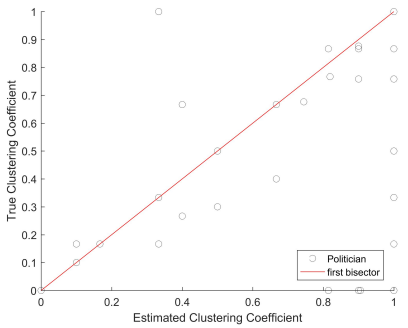

(a) Betweenness



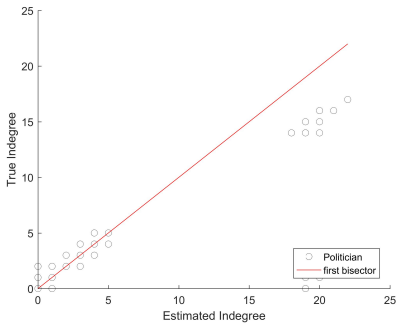

(b) Eigenvalue



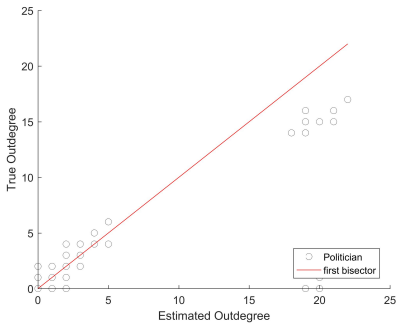

(c) Closeness




(d) Clustering




(e) Indegree




(f) Outdegree

NOTE. X-axis: estimated value of node-level statistic. Y-axis: true value of node-level statistic. The true network is generated using equations (26)-(27) and data from the alumni network. The connections of 200 politicians from the 111th Congress extracted at random are considered. The 111th Congress is randomly selected. The DGP is described in detail in Section 8.5. The estimated network is derived using the parameter estimates at the last iteration of the MCMC. The first of $\bar{c} = 5$ networks is considered. See Newman [2010] for the definition of network measures.

As further evidence of the good performance of the method, in Figure 2 we plot the values of individual network statistics for each node (betweenness, eigenvalue, closeness, clustering, and in- and out-degree).[32] Each point corresponds to a node of the network, with the estimated network statistic on the X-axis, the real 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, Table 1 shows aggregate network statistics in the simulated and true networks (density, assortativity, diameter, average distance, closeness, betweenness, degree, and clustering). Perhaps unsurprising at this stage, the table confirms that the estimated values are quite similar to the real ones.

**The circular network** Our second benchmark is a simulation in which $h_{i,j}$ is a "circular network" in which $i$ is linked to agents $i+j$ for $j \leq z_i$, where $z_i$ is an independent realization from a uniform distribution $U(0, \bar{z})$ for each legislator $i$ with $\bar{z} = 8$.[33], [34] While this setup is less realistic than the previous, it will prove useful in the next subsection for the comparative statics exercise because it allows us to easily change its features (size, sparsity, etc.). As in the previous section, here too we generate a network $G$ and $r = 5$ observations with $n = 200$ legislators in each. A similar analysis as in the previous section is presented in Figures 3 and 4. Results are very similar and lead to the same conclusions.

### 4.3.2 Sensitivity Analysis

Using the "circular network" discussed above as a benchmark, we now explore the performance of our method as we change key features of the environment, namely: the number of legislators $n$; the number of observations over time $r$; the sparsity of the network, as measured by $\bar{z}$; the elasticity of link formation $\phi$; and other important aspects of the network topological structure.

**Network size** We have simulated networks for $n = 25, 50, 100, 200,$ and 400 nodes. Figure 5 shows a box plot where each box represents the difference between the real and the estimated parameter.[35]

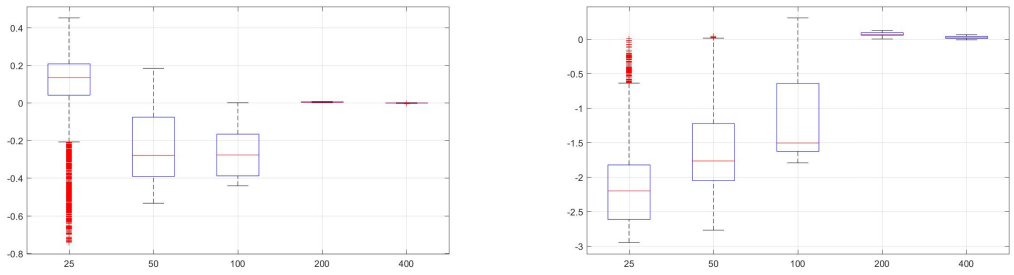
[32] See Newman [2010] for a formal definition of these centrality measures.

[33] We label this experiment the circular network because when $z_i = 1$ for all $i$, the resulting network is circular. The model we use is indeed a generalization of a simple circle that allows for a random number of connections.

[34] The variable $\bar{z}$ defines network sparsity. We have also repeated our simulation experiment using other distributions for $z_i$ (different from uniform). The results are not sensitive to different specifications. We do not report these further results for brevity. They remain available upon request. We will show the sensitivity of our methodology to network sparsity in Figure 10 and 11.

[35] The boxplot is constructed using the ergodic distribution generated by the Markov chain after a burning period of 10,000 iterations. We normalize the realizations by subtracting the true values from each of them, thus centering the distributions of all parameters at zero.

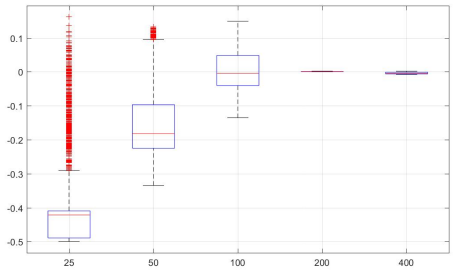
Figure 5: ESTIMATION BIAS FOR DIFFERENT NETWORK SIZES
- PARAMETERS -



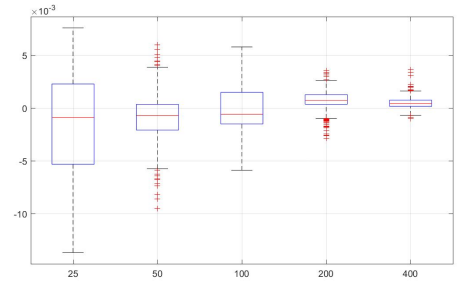

(a) $\varphi$




(b) $\phi$




(c) $\alpha$




(d) $\beta$

NOTE. X-axis: network size. Y-axis: distribution of the differences between estimated and true values in the MCMC after a burning period of 10,000 iterations. The estimated values are taken from the posterior distribution. The true values are fixed and generated using equations (26)-(27). The connections are generated from a circular network, as defined in Section 4.3.1. 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.

The graphs 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 $n = 200$ is already sufficient to have distances highly concentrated around zero. Figure 6 presents a boxplot where each box represents the difference between the real and the estimated node-level statistics. Here too the graphs show that the distributions concentrate around zero as $n$ increases.[36] Figure 7 shows the ROC curves

[36]We use the last simulated network to construct the measures. Given that the posterior of estimated parameters is highly concentrated around the true value, this network is very close to one constructed using estimates of each parameter and it is faster to compute. The centralities are normalized by subtracting their value in the true network. See Newman [2010] for the definition of the network centrality measures.

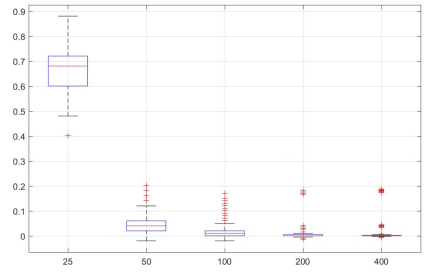
for each network size considered. The picture reveals a very high probability of detection of true links and a low probability of false positives for network sizes as small as 50 nodes. This simulation exercise provides ample evidence on the ability of our methodology to estimate social interactions in environments similar to the environment of our empirical investigation, where we will deal with five Congresses of about 400 politicians.

**Number of observations** We now vary the number of observations $r$ from 1 to 10 while keeping the number of nodes in the network constant at 200. Figures 8 and 9 have the same structure as Figures 5 and 6. They report the distribution of the estimation bias along the Markov chain in the estimated parameters, and the node-level characteristics when $r$ varies. 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 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.

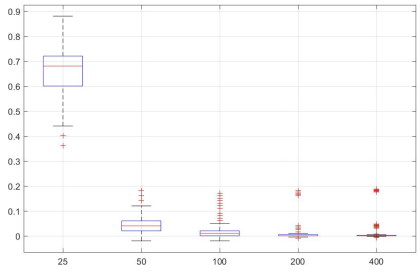
**Network sparsity** An important feature in the estimation of network models is the sparseness of connections between nodes (see De Paula et al. [2018] and Manresa [2016]). As hinted before, our simulation framework allows us to change network sparsity by varying the density of $Z$. A simple way to do this is to change $\bar{z}$, the maximum number of links that a node can have. In our benchmark model, $\bar{z}$ is equal to 8. We now set it to 4, 20, 40 and 100, while keeping constant all the other parameters.[37] As before, for this simulation exercise too we report the bias in the estimation of the parameters and of node-level statistics in Figures 10 and 11. These figures show that network sparsity is not a necessary condition for the estimation of our model, since the concentration of bias around zero does not appear to be related to network density.

[37] Given that $z_i$ is drawn from a uniform distribution, increasing the maximum by one unit implies a shift of the mean of one half. We have also repeated our simulation experiment using other distributions for $z_i$ (different from uniform). The results are not sensitive to different specifications. We do not report these further results for brevity. They remain available upon request.

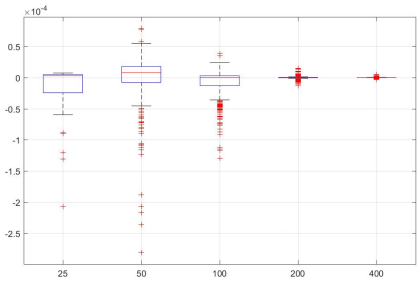
Figure 6: ESTIMATION BIAS FOR DIFFERENT NETWORK SIZES
- NODE-LEVEL STATISTICS -



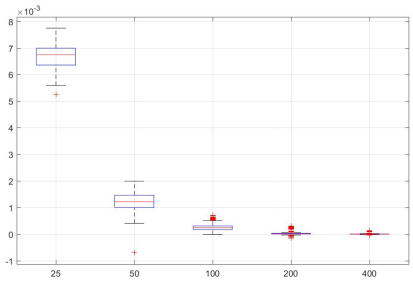

(a) Indegree



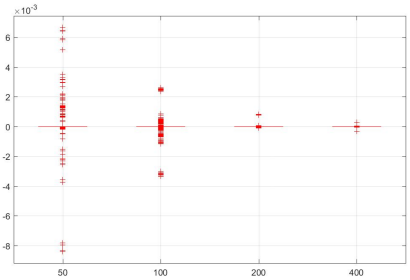

(b) Outdegree



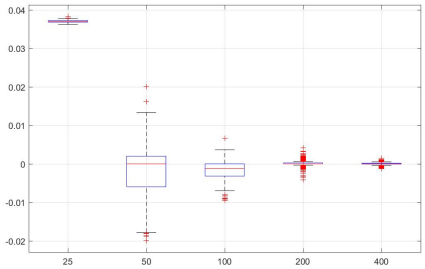

(c) Betweenness




(d) Closeness




(e) Eigenvector




(f) Clustering

NOTE. X-axis: network size. Y-axis: distribution of the differences between estimated and true values. See Newman [2010] for the definition of network measures. The true network is generated using equations (26)-(27). The connections are generated from a circular network, as defined in Section 4.3.1. The DGP is described in detail in Section 8.5. The estimated network is derived using the parameter estimates at the last iteration of the MCMC. Eigenvector centrality is not reported for the 25 nodes sample because the adjacency matrix has degenerate eigenvalues (i.e. with multiplicity greater than one). 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.

**The elasticity of network formation** A key parameter of our theoretical model is $\phi$, which captures the cost of linking among agents. When $\phi = 0$ the elasticity of network formation is zero and so model (10) is linear in $\theta_{i,j}$, as in standard spatial autoregressive models if $\theta_{i,j}$ is assumed to be the exogenous network. When $\phi > 0$, 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 $\phi$ changes. We set $\phi = 1, 2, 3$, and 4. Figure 12 and Figure 13 present, respectively, the distribution of the estimation bias for the parameters and node-level network statistics for each value of $\phi$. The boxplots reveal no systematic pattern across values of $\phi$, and that the distributions are mainly concentrated around zero for all values of $\phi$, with similar dispersion. These results thus indicate that performance of our methodology does not hinge on a particular value of $\phi$.

**Network topology** To conclude, we show the performance of our methodology under different network topologies. Figure 14 displays the true adjacency matrix, the associated estimated matrix using our methodology, and their difference when the true network is simulated under different rules. We consider a high density network (Panel a)), a low density network (Panel b)), a network structure coming from a highly nonlinear model (Panel c)), a reverse circular network (Panel d)), and the alumni network (Panel (e)).[38] Remarkably, the estimated linking structure closely follows the changes of the true network.

# 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, 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. Data are available online from the Legislative Effectiveness Project (http://www.thelawmakers.org).[39] We

[38] The high, medium and low density networks are circular networks with respectively $\overline{z} = 200$, $\overline{z} = 20$ and $\overline{z} = 8$. The network structure with the highly nonlinear model is the circular network with $\overline{z} = 8$ and $\phi = 4$. The reverse circular model is like the circular model but now $i$ is linked to agents $i - j$ for $j \leq z_i$, where $z_i$ is an independent realization from a uniform distribution $U(0, \bar{z})$ for each legislator $i$ with $\bar{z} = 8$. The alumni network is described in Section 4.1.

[39] Volden et al. [2013] have used this data to explore the legislative effectiveness of women for the 93rd - 110th Congresses. A similar index, Health ILESs, was proposed by Volden and Wiseman [2011] to examine which House

use information from five recent election cycles: the 109th Congress (election cycle 2004) to the 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. The Legislative Effectiveness Project thus provides data on the observed legislators' characteristics that are theoretically important for lawmaking effectiveness. They identify nine factors that are important for legislative effectiveness. We briefly discuss them here because we include them as controls in our analysis.

The first one is the number of years served as a member of the 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 the state legislatures may be more effective than legislators without similar experiences.[40] 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 are naturally associated with greater legislative effectiveness.[41]

The sixth factor captures *ideological considerations*. The Legislative Effectiveness Project data is merged with the http://voteview.com project data. It 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 captures the *demographic characteristics* of under-represented members in Congress. The experiences of women and ethnic minority legislators 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]).

members have been most successful at advancing health care bills for the 93rd to 110th Congresses.

[40] Previous legislative experience is captured using a dummy taking value one whether a legislator has previously served in their state legislature, and zero otherwise, and its interaction with the state's level of professionalism as measured by the index constructed by Squire [1992].

[41] This is true in general, with some exceptions. Volden and Wiseman [2014] document that the high level of effort required by the members of these committees results in a number of endorsed bills which is lower than that of the average House member, thus making the relationship with their LES scores negative. Minority party leaders are less likely to have their bills pass the House relative to other members of their party.

The eighth factor is 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). 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 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]).

Our analysis considers all of the legislator characteristics indicated in the Legislative Effectiveness Project. The control set $X_r$ in model (11) includes the number of years spent in Congress and its squared term, margin of victory and its squared term, dw-ideology, the size of the state congressional delegation, party, chairmanship, majority and minority party leadership, whether the representative is a member of the majority party, whether the representative is a member of the most powerful committees, previous legislative experience, gender, and race. Basic member characteristics, (party, gender, ethnicity, and seniority) are also included in the network formation model (12)-(13).[42] 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).[43] In our baseline version, 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]. Table A.1 in the Appendix provides a detailed description of

[42]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.

[43]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.

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 describe the estimation of model (11)-(13) using data from the 109th Congress (election cycle 2004) to the 113th Congress (election cycle 2012). Table 2 and Table 3 present the median value of the posterior distributions of the key parameters. Table 2 focuses on $\varphi$, $\alpha$, and $\phi$, and the $\beta_i$ contributing to $\varepsilon_i$. Column (1) of Table 3 shows the median values of the parameters governing the network formation: $\psi$, $\gamma$, and $\mu$. 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 and the convergence of the respective Markov chains are presented in Figures A1-A4 in the supplementary 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: $\varphi$, $\alpha$, and $\phi$. All of these parameters are found to be positive and statistically significant. The positive and statistically significant value of $\varphi$ measures the social multiplier in the model. The presence of social spillovers is further confirmed by a significantly positive $\alpha$, which measures the elasticity of $E_i = E_i - \varepsilon_i$ with respect to social connectedness and effort (social spillovers would be irrelevant if $\alpha$ were zero). A 1% increase in the social connectedness of $i$, as measured by $s_i$ in (1), induces a 0.74% increase in net effectiveness $E_i$. Regarding $\phi$, recall that $\phi = \epsilon_{g_{i,j},E_j}$, where $\epsilon_{g_{i,j},E_j}$ is the elasticity of $g_{i,j}$ with respect to $E_j$. When $\phi = 0$, the analysis reduces to that which simply assumes a spatial autoregressive model in which social connections are approximated by the alumni network. The fact that $\phi$ is significantly greater than zero highlights the importance of the fact that the network is endogenous: ceteris paribus, an increase in the effectiveness of $j$ by 1% leads to a median increase in the connection of another legislator $i$ to $j$ (i.e. $g_{i,j}$) of 0.07%.

Table 2 also presents the estimates of the legislators' individual characteristics directly affecting their effectiveness, i.e. the $\beta_i$s in $\varepsilon_i$s. The effects are in line with the existing theories of congressional politics, discussed in the previous section. Specifically, consistent with the idea that Congress members in positions of leadership are more effective, we find a significantly positive effect of being Chair and being Leader. Having served in State legislatures with higher levels of professionalism is also associated with greater legislative effectiveness in Congress, in line with the idea that relevant

skills in politics are acquired through professional experience. For seniority and the margin of victory, we detect nonlinear effects. For seniority, we find a convex effect with diminishing returns for low levels of seniority and positive returns for higher levels. 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 that freshmen have a lower effectiveness than the average Congress member (Volden and Wiseman, [2018]). The impact of the margin of victory is positive but 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 totally 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. Being a woman is associated with lower effectiveness, suggesting that the more integrative, collaborative, and consensual female approach to lawmaking may hinder womens' effectiveness (Jeydel and Taylor, [2003]). Non-white representatives, on the other hand, appear associated with a higher LES score, consistent with the idea that they are better organized in advocating policies, counting on natural coalitions of members in support of their legislation through groups such as the congressional Black Caucus (Volden and Wiseman, [2014]).

The findings discussed above should be contrasted with two benchmarks. First, the OLS estimate in which we ignore network effects. This is the estimate typical of the previous literature that has focused on the individual characteristics determining effectiveness while ignoring social spillovers. Second, a similar analysis imposing no social spillovers, but using the ABC methodology. This second estimation is interesting mostly as a robustness check because it should (and indeed it does) give us the same qualitative results as the OLS analysis. Table 4 collects the estimates. The first two columns report the results, respectively, from a standard OLS and the ABC estimator for the model (11) without social interactions, that is, when $\varphi$ is constrained to be equal to zero (and hence the effects of $\phi$ and $\alpha$ are also zero). The last column reports the estimates from Table 2, when network effects are allowed. Observe that when $\varphi$, $\phi$, and $\alpha$ are greater than zero, the marginal effect of the $k$-th covariate in model (11) is not just $\beta_k$, because $\beta_k$ directly affects $E_k$ which in turn affects all the remaining $E_{-k}$, which may affect $E_k$. An implication of this is that while the OLS model leads to a common estimate for the effect of a covariate, in the model with endogenous nonlinear network effects, the marginal effects are necessarily heterogeneous across individuals because they depend on the individual's position in the network. Another implication is that the $\beta_k$s should be expected to underestimate the true effect of the associated variable when we consider the social multiplier effect (see, e.g. LeSage, [2014] and Liu et al. [2017]).

We can formally test whether the model fit improves with the addition of network effects relative to a linear regression in which effectiveness is explained just by using the legislators' individual

characteristics (i.e $\varepsilon_i$). To this goal, note that the linear regression model is nested in our model for $\varphi = 0$, we can therefore use a standard partial F-test.[44] Results are presented in Column (3) of Table 4. The F-test rejects the hypothesis that the model with $\varphi \neq 0$ does not provide a significantly better fit than the model with $\varphi = 0$ (F value 29.111, $p < 0.0000$).

**The social network.** Column (1) of Table 3 provides an insight on the characteristics that matter for social connectedness in the U.S. Congress.[45] As apparent from Table 3, we find that belonging to the same party has the highest impact on the cost of forming a link. Remarkably, however, the effect of having a link in the alumni network is about one-fourth 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).

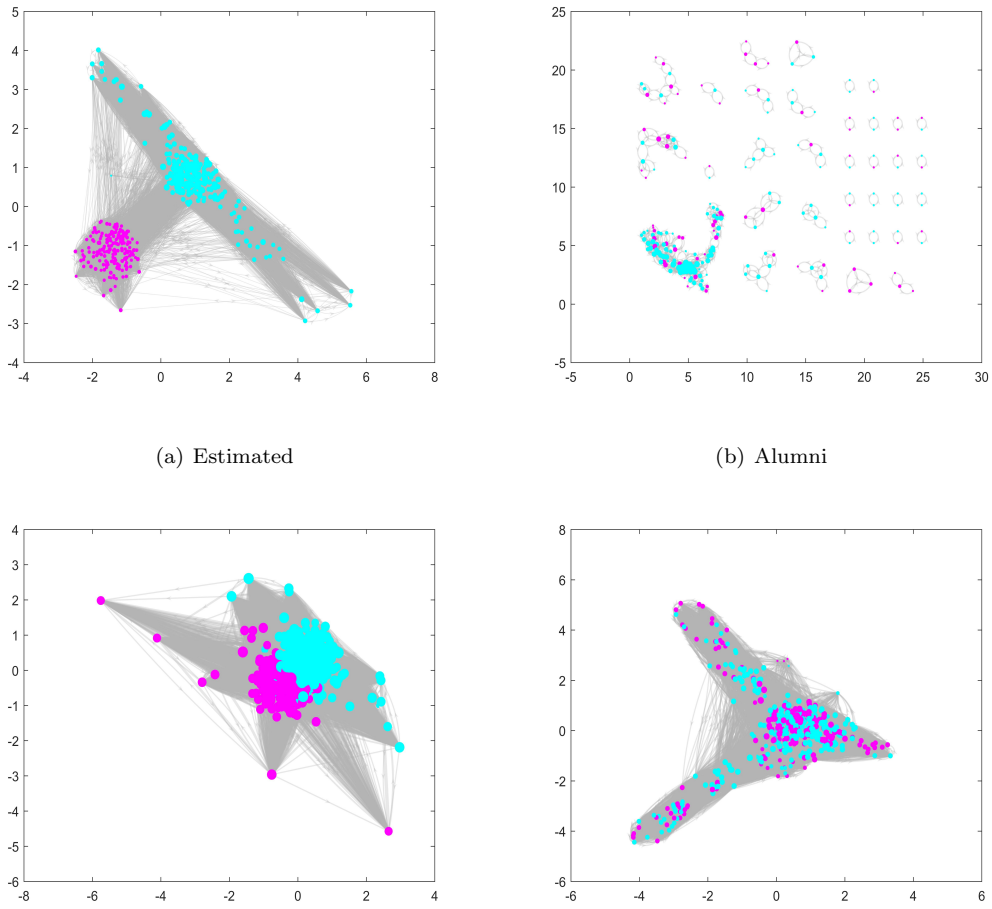
Using the estimates of Table 3, we can simulate the unobserved social network and use the simulation to study its characteristics. Figure 15 depicts the estimated network for the 111th Congress, simulated using the coefficients of Tables 2 and 3.[46] Three features of the social network uncovered by 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). It is clear that Democrats have a 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 as proxies for the true social network other observable adjacency matrices commonly used to study congressional behavior (such as the alumni network or the cosponsorship network) that do not exhibit such polarization. Figure 16 compares the estimated network, the alumni network, the committee network, and the cosponsorship network when coloring nodes by party affiliation. The picture reveals that the estimated network has a much better ability to capture polarization compared to the other networks.

[44] Let $RRS_1$ define the residual sum of squares of the unrestricted model [Column (3)] and $p_1$ the number of parameters. Let $RRS_2$ the residual sum of squares of the restricted model [Column (2)], 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.

[45] 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 6.4.

[46] We use force-directed layout (Fruchterman and Reingold [1991]), which pushes away from the cloud's center nodes with many links that are not connected among them. Qualitatively similar results are obtained with other Congresses, we omit the analysis of these cases here for brevity.

Figure 16: ESTIMATED VS OBSERVED NETWORKS



(a) Estimated

(b) Alumni

(c) Cosponsorship

(d) Committee

NOTE. The estimated network is derived using the parameter estimates at the last iteration of the MCMC for the 111th Congress. The color represents the party of the politician. Pink nodes are Republicans. The networks are represented with force-directed layout with five iterations (Fruchterman and Reingold [1991]). 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 alumni network is defined in Section 4.1. The $ij_{th}$ element of the committee network entry 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.

For the second feature, consider Panel (b) of Figure 15, 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 that we have just highlighted, legislators also develop weaker interparty connections. This is interesting because it confirms a theory that developing "weak ties" with an heterogeneous set of agents (especially, in the case of Congress, with different political ideologies) is essential to advance bills through the legislative process (see Granovetter [1973] and especially, for the U.S. Congress, Kirkland [2011]). An advantage of our approach is that it does not only provide a binary measure of connection, but also the intensive margin, which allows us to identify the strength of interactions.

The third feature we would like to highlight is an asymmetry between Democrats and Republicans: the Democratic cloud is more dispersed in terms of the node's connectivity than the Republican cloud, which features many members with few connections between each others. Democrats are also consistently more effective. This asymmetry reflects the fact that Democrats were the majority party in the 111th Congress, which clearly explains the higher average level of effectiveness.[47]

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 common unobserved shocks affect the legislators' effectiveness. This, however, can not be the case because the entries of the adjacency matrix are not estimated by relying exclusively on pairwise covariances between legislators' effectiveness. In the baseline model of Section 4.1, they are also functions of observable characteristics, orthogonal to unobserved shocks. As a result, an unobserved common shock hitting, say, politician $i$'s and $j$'s effectiveness does not affect the likelihood of estimating the link between them. It may do so only if it affects both the effectiveness and the likelihood of forming a link between legislators. In section 6.4 we extend our model to control for this issue, showing that the estimates are not affected by this type of concern.

# 6 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 $\frac{\phi}{(1+\phi)}\left(\frac{g_{i,j}}{\theta_{i,j}}\right)^{1+\frac{1}{\phi}}$. 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

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

$j$ versus $i$, denoted $\xi_{j,i}$. The cost of effort remains $\frac{\phi}{(1+\phi)}\left(\frac{\xi_{j,i}}{\theta_{i,j}}\right)^{1+\frac{1}{\phi}}$, the link is now:

$$g_{i,j} = \xi_{i,j}^{\vartheta} \cdot \xi_{j,i}^{1-\vartheta} \tag{14}$$

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 $i$ can not establish a link with $j$.

Under (14), $i$'s problem finds $\xi^{\mathbf{i}}$ solving:

$$\max_{\xi^{\mathbf{i}}} \left\{ \sum\nolimits_{j=1}^{n} \left[ \alpha\delta \cdot \xi_{i,j}^{\vartheta} \cdot \xi_{j,i}^{1-\vartheta} \cdot E_j(G,\varepsilon) - \frac{\phi}{(1+\phi)} \left(\frac{\xi_{i,j}}{\theta_{i,j}}\right)^{1+\frac{1}{\phi}} \right] \right\}. \tag{15}$$

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

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

which gives us:

$$\frac{\xi_{i,j}}{\xi_{j,i}} = \left(\frac{E_j}{E_i}\right)^{\frac{1}{2(1-\vartheta)+1/\phi}} \tag{16}$$

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

And therefore:

$$g_{i,j} = (\alpha\vartheta\delta)^{\phi} \theta_{i,j}^{1+\phi} \cdot \left[ (E_j)^{\phi} \left(\frac{E_i}{E_j}\right)^{\frac{(1-\vartheta)\phi}{2(1-\vartheta)+1/\phi}} \right]^{\vartheta} \left[ (E_i)^{\phi} \left(\frac{E_j}{E_i}\right)^{\frac{(1-\vartheta)\phi}{2(1-\vartheta)+1/\phi}} \right]^{1-\vartheta} \tag{17}$$

Combining the solution of (14) with (5), 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:*

$$E_i^* = \delta \cdot \sum\nolimits_{l \in \mathcal{N}} \left( (\alpha\vartheta\delta)^{\phi} \theta_{i,j}^{1+\phi} \cdot \left[ (E_l^*)^{\phi} \left(\frac{E_i^*}{E_l^*}\right)^{\frac{(1-\vartheta)\phi}{2(1-\vartheta)+1/\phi}} \right]^{\vartheta} \left[ (E_i^*)^{\phi} \left(\frac{E_l}{E_i}\right)^{\frac{(1-\vartheta)\phi}{2(1-\vartheta)+1/\phi}} \right]^{1-\vartheta} E_l^* \right) + \varepsilon_i \tag{18}$$

*for $i, j \in N$ and $G^* = (g^*_{i,j})_{i,j \in \mathcal{N}}$ is given by (17).*

Note that for $\vartheta = 1$, this expression is identical to (10). An advantage of (18) is that it would allow 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 need to assume some specific functional forms for the "production function" of effectiveness and the cost functions. From a theoretical point of view, most of the choices made 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:

$$E_i = \varphi A_i \cdot (s_i)^{\alpha} (l_i)^{1-\alpha} + \varepsilon_i \tag{19}$$

Under (19), legislator $i$'s characteristics do not only affect $E_i$ additively, through $\varepsilon_i$, they also affect it multiplicatively through $A_i$. In practice, in (19) we are allowing the spillover effect $\varphi$ to depend on the legislators' characteristics, i.e. $\varphi_i = \varphi \cdot A_i$.

Following the same steps as in Section 3, where we incorporate the optimal level of effort in (19), we obtain that in stage 2, when $G$ is given, effectiveness are given by:[48]

$$\mathbf{E} = (I - \delta \cdot \Gamma(A, \alpha) \cdot G)^{-1} \varepsilon \tag{20}$$

where $\Gamma(A, \alpha)$ is a diagonal matrix with $i$th diagonal element equal to $A_i^{\frac{1}{\alpha}}$. Condition (20) generalizes (5) by avoiding to impose that the social spillover is the same for all legislators: some may benefit more or less than others. In (20) the social spillover depends on the legislators characteristics affecting $A_i$. This generalization gives us additional flexibility in 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 $\varsigma_2$ a fixed effect.[49] Solving for the optimal social connections in the first stage as in Section 3.3, in an interior solution we have:

$$E_i = \delta \cdot [\alpha\delta]^{\phi} \left[(\mathbf{X}_{i,r})' \varsigma_1 + \varsigma_2\right]^{\frac{1+\phi}{\alpha}} \cdot \sum\nolimits_j \left[\theta_{i,j} E_j\right]^{1+\phi} + \varepsilon_i \tag{21}$$

[48] In (20) 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.

[49] Naturally, we can include also random terms, as $\varepsilon_i$ in Section 4.1.

with $g_{i,j} = \left[\alpha\delta \cdot A_i^{\frac{1}{\alpha}}\right]^{\phi} \cdot \theta_{i,j}^{1+\phi} E_j^{\phi}$. We can then use (21) in step C2 of our algorithm exactly as described in Section 4.2.

### 6.3 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 $j$ can not hurt $i$.[50] 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.

### 6.4 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 $E_{i,r}$s. Suppose that an unobserved characteristic of node $i$ in network $r$, $\eta_{r,i}$, matters for the network formation process. Then we can extend model (13) in the following way:[51]

$$\lambda_{i,j,r} = \iota + \gamma h_{i,j,r} + \sum\nolimits_l g(X_{r,li}, X_{r,lj})\psi_l + m(\eta_{r,i}, \eta_{r,j})\kappa, \tag{22}$$

where $m(\cdot,\cdot)$ is a distance function and $\kappa$ is a parameter. We allow the outcome error term $\epsilon_{r,i}$ to be correlated with $\eta_{r,i}$, with $\epsilon_{r,i} = \sum_{l=1}^{L} \mu_l \eta_{r,i}^l + u_{r,i}$.[52] In any step of the ABC algorithm (described

[50] 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 $i$.

[51] The model can be further extended to a vector of unobservable characteristics and to multiple observed adjacency matrices.

[52] 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 $\sigma_\eta$ is a threshold that we set equal to the variance of $\eta$. 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 $\sigma_u$. 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 higher individual propensity to link (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.

in Section 8.6), $\epsilon_{r,i}$ is thus generated as a function of $\eta_{r,i}$ and $\mu = (\mu_1, ..., \mu_L)$, which are estimated as the rest of the model's parameters.

The results associated to this extension of the empirical model are shown in column 2 of Table 3 for the network formation model (equation (22)) and in Table A.2 in the supplementary online Appendix for the outcome equation (equation (11) augmented with $\epsilon_{r,i} = \sum_{l=1}^{L} \mu_l \eta_{r,i}^l + u_{r,i}$). In the first column of Table A2, we report our baseline estimates of Table 4, Column (3), for comparison. The evidence on the positive and statistically significant estimates of $\phi$, $\varphi$ and $\alpha$ holds true.

By allowing for unobserved variables that are correlated with the legislators' effectiveness to affect the choice of social links, our approach is along the lines of Hsieh and Lee [2014], Boucher et al. [2019], and Goldsmith-Pinkham and Imbens [2013]. Our analysis, however, differs in two significant ways. First, in our model the true network is unobservable–only the fundamental variables determining the network formation are observable. Second, and most importantly, the network endogeneity is explicitly modeled. This is what gives us the nonlinear representation in (8).

# 7 Conclusions

We have presented a model of endogenous formation of social connections in the U.S. Congress to study the determinants of legislators' effectiveness. In the model, legislators first invest in their social connections with other legislators, then they exert effort to achieve their favorite legislative projects, taking the connections as given. The ability of legislators to achieve their goals in Congress–their legislative effectiveness–depends on the social connections that they have previously established, the effectiveness of the other legislators with whom they are socially connected, their efforts, and their characteristics. Using data from the 109th to 113th U.S. Congresses, we have structurally estimated the model to gain insight on the legislators social network and its effect on the legislators effectiveness.

Two methodological contributions are at the core of our analysis. First, we introduce a new equilibrium concept that we call the Competitive Network Equilibrium. Drawing an analogy with the approach in competitive analysis to study "price taker" 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 in the estimation technique, which is based on approximate computation methods (ABC), an approach to draw Bayesian inference without an analytic likelihood function.

Our analysis suggests that social connections in the U.S. Congress are an important factor in

determining legislative effectiveness, thus confirming an idea that has been long informally discussed in journalistic and historical accounts of how Congress works. Perhaps more importantly, the estimation of the social network gives us an insight into the determinants of social connections. Not surprisingly, 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 that alumni connections are also a key factor, about one-fourth as large as party affiliation. We also find that weak links across the political lines are important, an observation long suggested in the sociological and political science literature (see Granovetter [1973] and more recently Kirkland [2011]). The model, moreover, allows us to estimate social network effects using only information on outcomes. This is an improvement over the existing practice, where some observable adjacency matrix is used as a proxy for social connections. As we have discussed in Section 5.2, by exclusively using alumni networks as a proxy, we ignore the importance of party affiliation, and by using only cosponsorship, we ignore the importance of alumni connections.

We have discussed in Section 6 a few extensions of the model and the estimation techniques. More generally, we believe the approach used in this work can more broadly be applied to study social networks in other 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.

# References


[1] Acemoglu D. and P. Azar (2018), "Endogenous production networks," Working paper.

[2] Austen-Smith, D., and J. Banks (1988): "Elections, Coalitions, and Legislative Outcomes," *American Political Science Review*, 82(2), 405-422.

[3] Badev, A. (2017): "Discrete Games in Endogenous Networks: Equilibria and Policy,"Available at arXiv:1705.03137v1.

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

[5] Baron, D., and D. Diermeier (2001): "Elections, Governments, and Parliaments Under Proportional Representation," *Quarterly Journal of Economics*, 116(3), 933-967.

[6] Baron, D., and J. Ferejohn (1989): "Bargaining in Legislatures," *American Political Science Review*, 83(4), 1181-1206.

[7] Battaglini, M. (2019): "Coalition Formation in Legislative Bargaining," Discussion paper, National Bureau of Economic Research.

[8] Battaglini, M., F.W. Crawford, E. Patacchini, and S. Peng (2018): "Inferring Networks from Outcomes," Working paper.

[9] Battaglini, M., V. Leone Sciabolazza, and E. Patacchini (2019): "Effectiveness of Connected Legislators," *American Journal of Political Science*, (forthcoming).

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

[11] Battaglini, M., and E. Patacchini (2019): "Social Networks in Policy Making," *Annual Review of Economics* (forthcoming).

[12] Bonacich, P. (1987): "Power and Centrality: A Family of Measures," *American Journal of Sociology*, 92(5), 1170-1182.

[13] Boucher, V. (2018): "Equilibrium Homophily in Networks," Working paper.

[14] Boucher, V., C.S. Hsieh and L.F. Lee (2019): "Specification and Estimation of Network Formation and Network Interaction Models with the Exponential Probability Distribution," Working paper.

[15] Breza, E., A.G. Chandrasekhar, T.H. McCormick, and M. Pan (2017): "Using Aggregated Relational Data to Feasibly Identify Network Structure Without Network Data," Discussion paper, National Bureau of Economic Research.

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

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

[18] Canen, N., M.O. Jackson, and F. Trebbi (2017): "Endogenous Networks and Legislative Activity," Available at SSRN 2823338.

[19] Christakis, N.A., J.H. Fowler, G.W. Imbens, and K. Kalyanaraman (2010): "An Empirical Model for Strategic Network Formation, " Discussion paper, National Bureau of Economic Research.

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

[21] De Paula, A. (2017): "Econometrics of Network Models," *Advances in Economics and Econometrics: Theory and Applications, Eleventh World Congress* (Econometric Society Monographs), Cambridge University Press.

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

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

[24] Eulau, H. (1962): "Comparative Political Analysis: A Methodological Note," *Midwest Journal of Political Science*, 6(4), 397-407.

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

[26] Florens, J.P., M. Mouchart and J.M. Rolin (1990): "Elements of Bayesian Statistics," Routledge.

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

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

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

[30] Fu, Y.X., and W.H. Li (1997): "Estimating the Age of the Common Ancestor of a Sample of DNA Sequences," *Molecular Biology and Evolution*, 14(2), 195-199.

[31] Glassman, M., and A. Wilhelm (2017): "Congressional Careers: Service Tenure and Patterns of Member Service, 1789-2017," Congressional Research Service Report n. 7-5700.

[32] 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.

[33] Graham, B. S. (2017) "An econometric model of network formation with degree heterogeneity," *Econometrica* 85 (4), 1033-1063.

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

[35] Handcock, M.S., and K.J. Gile (2010): "Modeling Social Networks from Sampled Data," *The Annals of Applied Statistics* 4(1), 5-25.

[36] Harmon, N., R. Fisman, and E. Kamenica (2017): "Peer Effects in Legislative Voting," *American Economic Journal: Applied Economics* (forthcoming).

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

[38] 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.

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

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

[41] 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.

[42] Kass, R.E., B. Carlin, A. Gelman, R. Neal (1998): "Markov chain Monte Carlo in practice: A roundtable discussion," *The American Statistician*, 52, 93-100.

[43] 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.

[44] Kelejian, H.H., and I.R. Prucha (2010): "Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances, " *Journal of Econometrics*, 157(1), 53-67.

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

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

[47] Koskinen, J.H., G. Robins, and P.Pattison (2010): "Analyzing Exponential Random Graph (p-star) Models with Missing Data Using Bayesian Data Augmentation," *Statistical Methodology*, 7(3), 366-384.

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

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

[50] LeSage, J.P. (2014): "What Regional Scientists Need to Know About Spatial Econometrics," *Review of Regional Studies*, 44(1), 13-32.

[51] Lindley D.V. (1971): "Bayesian statistics: a review" Philadelphia, SIAM.

[52] Liu, X., E. Patacchini, Y. Zenou, and L. Lee (2012): "Criminal Networks: Who is the Key Player?" CEPR Discussion Paper No. 8772.

[53] Liu, X., E. Patacchini, and E. Rainone (2017): "Peer Effects in Bedtime Decisions Among Adolescents: A Social Network Model with Sampled Data," *The Econometrics Journal*, 20(3), S103-S125.

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

[55] 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.

[56] Masket, S. (2008): "Where You Sit Is Where You Stand: The Impact of Seating Proximity on Legislative Cue-Taking," *Quarterly Journal of Political Science*, 3(3), 301-311.

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

[61] Morelli, M. (1999): "Demand Competition and Policy Compromise in Legislative Bargaining," *American Political Science Review*, 93(4), 809-820.

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

[63] Oberfield, E. (2018): "A Theory of Input-Output Architecture," *Econometrica*, 86, 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] Parigi, P., and L. Sartori (2014): "The Political Party as a Network of Cleavages: Disclosing the Inner Structure of Italian Political Parties in the Seventies," *Social Networks*, 36(1), 54-65.

[66] Peoples, C.D. (2008): "Interlegislator Relations and Policy Making: A Sociological Study of Roll-Call Voting in a State Legislature," *Sociological Forum*, 23(3), 455-480.

[67] Plagnol, V., and S. Tavare (2004): in *Proceedings of the 5th International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, ed. Niederreiter, H. (Springer, Heidelberg).

[68] Porter, M.A., P.J. Mucha, M.E.J. Newman, and C.M. Warmbrand (2005): "A Network Analysis of Committees in the U.S. House of Representatives," *Proceedings of the National Academy of Sciences*, 102 (20), 7057-7062.

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

[70] Rice, S.A. (1927): "The Identification of Blocs in Small Political Bodies," *American Political Science Review*, 21(3), 619-627.

[71] Rice, S.A. (1928): "Quantitative Methods in Politics," *Journal of the American Statistical Association*, 33(201), 126-130.

[72] Robert C.P., J. Cornuet, J. Marin, and N.S. Pillai (2011): "Lack of Confidence in Approximate Bayesian Computation Model Choice," *Proceedings of the National Academy of Sciences*, 108(37): 15112-15117.

[73] 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.

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

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

[76] Rogowski, J.C., and B. Sinclair (2012): "Estimating the Causal Effects of Social Interaction with Endogenous Networks," *Political Analysis*, 20(3), 316-328.

[77] Routt, G.C. (1938): "Interpersonal Relationships and the Legislative Process," *The Annals of the American Academy of Political and Social Science*, 195(1), 129-136.

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

[79] Seidmann D., E. Winter, and E. Pavlov (2007): "The Formateurs' Role in Government Formation," *Economic Theory*, 31(3), 427-445.

[80] Shalizi, C.R., and A. Rinaldo (2013): "Consistency Under Sampling of Exponential Random Graph Models," *The Annals of Statistics*, 41(2), 508-535.

[81] Sheng, S. (2018): "A Structural Econometric Analysis of Network Formation," Working paper.

[82] Shepsle, K.A., and B.R. Weingast (1981): "Political Preferences for the Pork Barrel: A Generalization." *American Journal of Political Science*, 25(1), 96-111.

[83] 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.

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

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

[86] Taschereau-Dumouchel, M. (2019), "Cascades and Fluctuations in an Economy with an Endogenous Production Network," Working paper.

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

[88] 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.

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

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

[91] 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.

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

[93] Weingast, B., and W. Marshall (1988): "The Industrial Organization of Congress; or, Why Legislatures, Like Firms, Are Not Organized as Markets," *Journal of Political Economy*, 96(1), 132-163.

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

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

# 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 (7) 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 (7) 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: $g^*_{i,j} \leq (\theta_{i,j})^{1+\phi} \left(\alpha \delta E^*_j\right)^{\phi}$, satisfied as an equality in an interior solution. Combining this with the necessary and sufficient condition (5) 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 $E^*_j$s and the $n(n-1)$ variables $g^*_{i,j}$ for $i \in N$ and $j \in N \backslash \{i\}$) described in (8) and (9) (given that the elements in the diagonal of $G^*$ are all zero).[53] ■

## 8.2 Proof of Proposition 2

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

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

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 it is 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

[53] 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 $i = 1, ..n$.

can establish a link. We can now write:

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

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

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

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

## 8.3 Proof of the result in Example 2

First, consider an equilibrium with no connections. A necessary and sufficient condition for its existence is that a legislator, expecting no connections by the other players, finds it optimal to establish no connections as well. In this equilibrium, the effectiveness of an agent $j$ is $\varepsilon$. 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 $\varepsilon \geq (1-2\delta\overline{g})\,/\,(\alpha\delta)$. ■

## 8.4 Proof of Proposition 3

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

$$F_{\nu}^{*}(\omega) = \int_x Q_{\nu}\left(x \to \omega\right) F_{\nu}^{*}(x)dx \tag{23}$$

where $Q_{\nu}\left(x \to \omega\right)$ is the transition probability of Algorithm $C$. We need to prove that $f\left(\omega\,|\mathcal{D}_{\nu}\right)$ is the unique fixed-point of (23). Let $\omega$ be the vector of parameters to estimate and let $\mathcal{D}_{\nu}$ be defined as in Section 4.2. We have:

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

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

$$\frac{\pi\left(\omega'\right) q(\omega' \to \omega)}{\pi\left(\omega\right) q(\omega \to \omega')} \leq 1 \tag{24}$$

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

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

where in the last stage we used the fact that if (24) holds, then $\frac{\pi\left(\omega'\right) q(\omega \to \omega')}{\pi(\omega) q(\omega' \to \omega)} > 1$ and thus $h(\omega', \omega) = 1$.

Assume now that (24) does not hold. We have for any $\omega \neq \omega'$:

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

From these two cases we conclude that:

$$f\left(\omega\,|\mathcal{D}_\nu\right) Q_\nu\left(\omega \to \omega'\right) = f\left(\omega'\,|\mathcal{D}_\nu\right) \cdot Q_\nu\left(\omega' \to \omega\right) \tag{25}$$

If we integrate both sides of (25) by $\omega'$ 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 (23) 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 (11):

$$E_i = \alpha^\phi \delta^{1+\phi} \sum_{j \neq i} (\theta_{i,j} E_j)^{1+\phi} + \beta X_i + \epsilon_i \tag{26}$$

Here $X_i$ is a unidimensional random variable generated from a normal distribution $N(0, \sigma_x)$ and $\theta_{i,j}$ is generated from (12) where:

$$\lambda_{i,j} = v(\iota + \gamma h_{i,j}). \tag{27}$$

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 baseline "circular model," we set $\bar{z} = n/w = 200/25 = 8$. This number is selected in order to have a degree close to the alumni network (about 0.04, see Table A.3). In Section 4.3.2 we also change $\bar{z}$ and show the insensitivity of our results to it. For both models, we set $\sigma_x = 1$, $\beta = 1$, $\alpha = 0.5$, $\phi = 2$, $\sigma_\epsilon = 0.1$ and $\varphi = 0.5$. The linking parameters are set as $\iota = -1$, $\gamma = 2$ and $v = 7$. The parameters $(\iota, \gamma, \psi_l, \alpha, \phi, \varphi)$ 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 diameter, the average distance of the politicians' cosponsorship network, and the betweenness of the committee network (in which two legislators are linked if they belong to the same committee). See Table 1 and Table A.3. All of the results discussed in Section 4.3 are robust to alternative configurations of the parameters.

## 8.6 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 (11)-(13):

$$\begin{aligned}
\phi &\sim TN_{\{0,\infty\}}(\phi_0, \Phi_0), & (\psi, \gamma, \iota) &\sim N_{K_l+2}(\omega_0, \Omega_0) \\
\alpha &\sim TN_{\{0,1\}}(\alpha_0, \Delta_0), & \varphi &\sim U[-\chi, \chi] \\
\eta_{i,r} &\sim N(0, \eta_0) & \sigma_\epsilon^2 &\sim TN_{\{0,\infty\}}(\sigma_0, \Sigma_0) \\
\beta &\sim N_K(\beta_0, B_0) & \zeta_r &\sim N(0, \sigma_\zeta) \\
\gamma &\sim N(\gamma_0, \Gamma_0) & \mu &\sim N(0, \mu_0)
\end{aligned}$$

where $U[\cdot]$ , $TN_{\{a,b\}}(\cdot)$ and $N(\cdot)$ are, respectively, the uniform, truncated normal (with $a$ and $b$ as lower and upper bounds), and normal distributions. For our key parameter of interest measuring the social externality, $\varphi$, we adopt a uniform (uninformative) prior centered at zero, as suggested in Smith and LeSage [2004] for spatial autoregressive models. Following Hsieh and Lee [2014], we adopt standard normal priors for the parameters of covariates in the outcome and link formation equations (i.e. $\beta$, $\psi, \iota$), and for $\eta_{i,r}$ and $\mu$ (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. As both the Cobb-Douglass coefficient $\alpha$ and the endogenous network parameter $\phi$ are constrained to be non-negative, we use truncated normal distributions. The truncated normal prior of $\alpha$ assumes values between zero and one. The truncated normal prior of $\phi$ is only constrained to be positive.

We set the hyperparamethers as follows. For the parameters for which we have no prior information, we chose neutral values: $\phi_0$ and $\alpha_0$ are set equal to zero; $\Phi_0$ and $\Delta_0$ are set equal to 0.5; $\eta_0$ (i.e., the variance of the prior for $\eta$) is set equal to one; $\sigma_0$ is set equal to zero; $\chi$ is set at 0.5; $\Sigma_0$ is a diagonal matrix with 0.5 on its diagonal elements; $\sigma_\zeta$ 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; $\beta_0$ 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 correspondent variance covariance matrix; $\omega_0$ is set equal to the logit point estimates obtained by regressing the pairwise controls on the cosponsorship network entries, and $\Omega_0$ is the correspondent variance covariance matrix. Hyperparameters in the prior distribution for the fixed effects $\zeta$s 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 Calvò-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 $\phi$, $\varphi$, $\alpha$, $\psi, \gamma, \iota$ and $\eta_{i,r}$ only if they satisfy Proposition 2, i.e. when $\delta < (1/\overline{\theta}) \cdot \left[1/\left((1+\phi)\,\alpha^{\phi}\overline{m}\right)\right]^{1/(1+\phi)}$. Observe that $\psi, \gamma, \iota$ 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 are reported in the supplementary online Appendix in Figures A.2-A.4.

**Sampling Algorithm** The initial state of the Markov chain

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

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.[54] To draw new values for each parameter ($\omega_i'$) at iteration $t$**,** we use a normal kernel, with mean equal to the current value and variance set at a parameter-specific tuning parameter $c$:

$$N(\omega_{i,t}, c). \tag{28}$$

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]).[55] 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

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

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

superior to 80% in drawing new value. Namely:

$$\begin{aligned} &\text{if } t_{A,i}/t \leq 0.2 && \text{then } c_{t+1} = c_t/1.005, \\ &\text{if } t_{A,i}/t \geq 0.8 && \text{then } c_{t+1} = c_t \times 1.005, \\ &\text{if } 0.2 \leq t_{A,i}/t \leq 0.8 && \text{then } c_{t+1} = c_t, \end{aligned} \tag{29}$$

where $t_{A,i}$ is the number of accepted draws at iteration $t$. The sequence $c_t$ converges after iteration 10,000 to a level $c_\infty$. In the second phase, the parameter is set at its convergence level $c_\infty$. This mechanism guarantees a bounded acceptance rate and convergence to optimal tuning. Figure A.6 in the supplementary online Appendix reports the acceptance rate ($t_{A,i}/t$), i.e. the probability of moving from $\omega_i$ to $\omega_i'$, for each of our parameters over the $MCMC$ iterations. We can see that rates converge to values ranging from 20 to 35 percent, showing good mixing properties.

Our algorithm relies on the choice of the tolerance $\epsilon$, i.e. 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 $\epsilon$ equal to the first quartile of the observed tolerance distribution. Specifically, for the first $10,000$ iterations we substitute step $C2$ with:

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

In this way the algorithm moves to regions of the parameter space where the distance from the real data is lower. Figure A.5 shows the rapid convergence of the distance between the simulated and the real data.[56] We use the Manhattan norm distance, $\rho(z(\mathbf{E},\omega)) = ||z(\mathbf{E},\omega)||_1 = \sum |z_i(\mathbf{E},\omega)|$. The results do not change significantly using different norms.

[56] 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.

# Tables

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

| | Estimated | True |
|---|---|---|
| Density | 0.0694 | 0.0583 |
| Assortativity | 26.7911 | 21.4879 |
| Diameter | 6.0000 | 6.0000 |
| Average distance | 2.8238 | 3.0403 |
| Closeness | 0.4066 | 0.3786 |
| Betweenness | 0.1206 | 0.0923 |
| Degree | 0.1684 | 0.1441 |
| Clustering | 0.7113 | 0.6558 |

NOTE. The true network is generated using equations (26)-(27). 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 measures.

Table 2: MAIN ESTIMATION RESULTS

| Dependent variable: legislative effectiveness | |
|---|---|
| | ABC Endogenous network effects |
| $\varphi$ | 0.0263 *** |
| | [1.0000] |
| $\alpha$ | 0.7807 *** |
| | [1.0000] |
| $\phi$ | 0.0741 *** |
| | [1.0000] |
| Party | -0.0116 *** |
| | [0.0000] |
| Gender | -0.0020 *** |
| | [0.0000] |
| Non white | 0.0042 *** |
| | [1.0000] |
| Seniority | -0.0015 *** |
| | [0.0000] |
| Seniorty$^2$ | 0.0002 *** |
| | [1.0000] |
| DW ideology | -0.0268 *** |
| | [0.0000] |
| Margin of victory | 0.0319 *** |
| | [1.0000] |
| Margin of victory$^2$ | -0.0319 *** |
| | [0.0000] |
| Committee chair | 0.1789 *** |
| | [1.0000] |
| Powerful committee | -0.0037 *** |
| | [0.0000] |
| Delegation size | -0.0001 |
| | [0.1240] |
| Leader | 0.0040 *** |
| | [1.0000] |
| State legislative experience | -0.0058 *** |
| | [0.0000] |
| State legislative experience * | 0.0300 *** |
| State legislative professionalism | [1.0000] |
| Congress fixed effects | Yes |
| Partial F test [no network effects] | 29.111 |
| p-value | 0.0000 |
| N. Obs. | 2,176 |

NOTE. Estimates of parameters in equation (11). 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.1. *, **, *** indicate statistical significance at the 10, 5 and 1 percent levels, based on empirical p-values.

Table 3: ESTIMATION RESULTS - LINK FORMATION

Dependent variable: probability of forming a link

| | ABC Endogenous network effects (1) | ABC Endogenous network effects with unobservables (2) |
|---|---|---|
| Alumni network | 0.9325 *** | 0.9652 *** |
| | [1.0000] | [1.0000] |
| Constant | -4.6279 *** | -4.9112 *** |
| | [0.0000] | [0.0000] |
| Seniority | 0.5206 *** | 0.4297 *** |
| | [1.0000] | [1.0000] |
| Leader | 0.0200 | -0.0387 ** |
| | [0.7474] | [0.0339] |
| Gender | -0.6530 *** | 0.1593 |
| | [0.0000] | [0.8849] |
| Non white | 1.1445 *** | 1.5384 *** |
| | [1.0000] | [1.0000] |
| Party | 4.2029 *** | 3.7434 *** |
| | [1.0000] | [1.0000] |
| Unobservables | | -0.0127 |
| | | [0.2209] |
| Congress fixed effects | Yes | Yes |
| N. Obs. | 2,176 | 2,176 |

NOTE. Estimates of parameters in equation (13). 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.1. *, **, *** indicate statistical significance at the 10, 5 and 1 percent levels, based on empirical p-values.

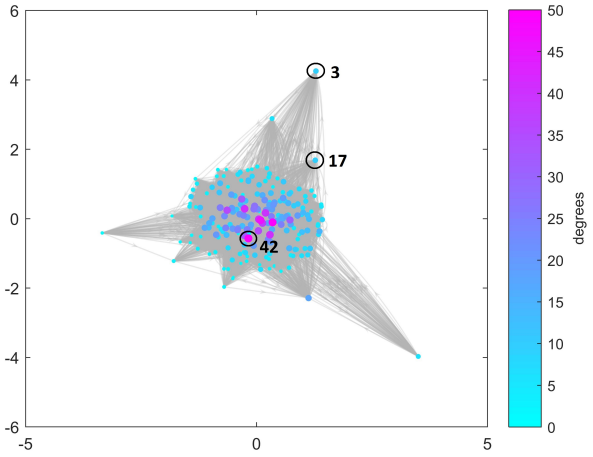
Table 4: MAIN ESTIMATION RESULTS
-WITH AND WITHOUT NETWORK EFFECTS-

| Dependent variable: Legislative Effectiveness | | | |
|---|---|---|---|
| | OLS | ABC<br>No network effects | ABC<br>Network effects |
| | (1) | (2) | (3) |
| $\varphi$ | - | - | 0.0263 ***<br>[1.0000] |
| $\alpha$ | - | - | 0.7807 ***<br>[1.0000] |
| $\phi$ | - | - | 0.0741 ***<br>[1.0000] |
| Party | -0.0173 ***<br>[0.0000] | -0.0171 ***<br>[0.0000] | -0.0116 ***<br>[0.0000] |
| Gender | 0.0013<br>[0.6338] | 0.0034<br>[0.7575] | -0.0020 ***<br>[0.0000] |
| Non white | -0.0026<br>[0.2830] | -0.0002<br>[0.4855] | 0.0042 ***<br>[1.0000] |
| Seniority | -0.0020 **<br>[0.0190] | -0.0027 ***<br>[0.0000] | -0.0015 ***<br>[0.0000] |
| Seniorty$^2$ | 0.0003 ***<br>[1.0000] | 0.0002 ***<br>[1.0000] | 0.0002 ***<br>[1.0000] |
| DW ideology | -0.0208 ***<br>[0.0080] | -0.0302 ***<br>[0.0000] | -0.0268 ***<br>[0.0000] |
| Margin of victory | 0.0383 **<br>[0.9726] | 0.0409 ***<br>[1.0000] | 0.0319 ***<br>[1.0000] |
| Margin of victory$^2$ | -0.0259 *<br>[0.0875] | -0.0457 ***<br>[0.0000] | -0.0319 ***<br>[0.0000] |
| Committee chair | 0.1822 ***<br>[1.0000] | 0.1790 ***<br>[1.0000] | 0.1789 ***<br>[1.0000] |
| Powerful committee | -0.0090 ***<br>[0.0040] | -0.0041<br>[0.1351] | -0.0037 ***<br>[0.0000] |
| Delegation size | -0.0002 *<br>[0.0601] | -0.0002 **<br>[0.0440] | -0.0001<br>[0.1240] |
| Leader | 0.0082<br>[0.8930] | 0.0067 ***<br>[1.0000] | 0.0040 ***<br>[1.0000] |
| State legislative experience | -0.0032<br>[0.2661] | -0.0018<br>[0.1255] | -0.0058 ***<br>[0.0000] |
| State legislative experience *<br>State legislative professionalism | 0.0248 **<br>[0.9508] | 0.0344 ***<br>[1.0000] | 0.0300 ***<br>[1.0000] |
| Congress fixed effects | Yes | Yes | Yes |
| Partial F test [no network effects] | | | 29.111 |
| p-value | | | 0.0000 |
| N. Obs. | 2,176 | 2,176 | 2,176 |

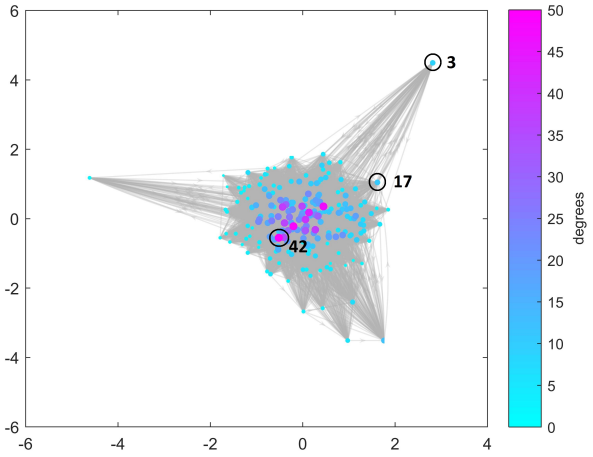
NOTE. Estimates of parameters in equation (11). ABC estimated coefficients are reported in columns (2)-(3). The median of the posterior distribution estimated with the ABC algorithm is reported for each coefficient. The empirical p-value of zero on the estimated posterior is reported in brackets. OLS estimated coefficients are reported in column (1). For comparison the p-value of zero on a normal distribution with mean equal to the OLS point estimate and variance equal to the estimated variance of each coefficient is reported in brackets. A precise definition of control variables can be found in Table A.1. *, **, *** indicate statistical significance at the 10, 5 and 1 percent levels, based on empirical p-values.

## Figures

Figure 3: CIRCULAR NETWORK ESTIMATION
- GOODNESS OF FIT -



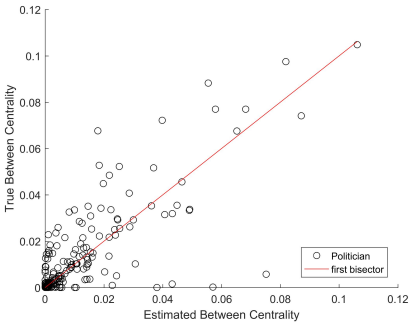

(a) Estimated



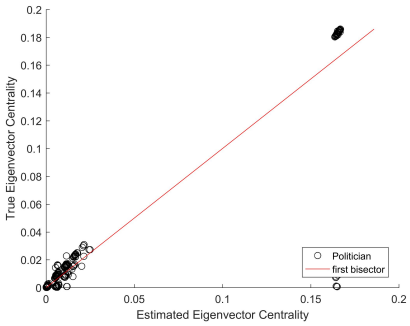

(b) True

NOTE. Panel (a) and (b) represent the estimated and true network respectively. The true network is generated using equations (26)-(27). The connections are generated from a circular network, as defined in Section 4.3.1. The DGP is described in detail in Section 8.5. The estimated network is derived using the parameter estimates at the last iteration of the MCMC. The first of $\bar{c} = 5$ networks is visually represented using the force-directed layout algorithm with five iterations. 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.

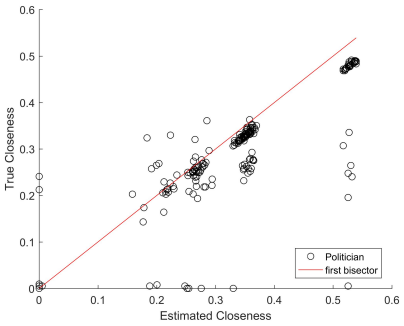
Figure 4: NODE-LEVEL STATISTICS
- ESTIMATED VS TRUE CIRCULAR NETWORK -



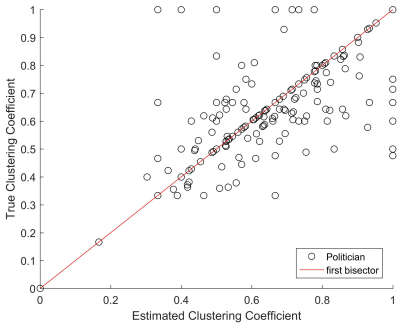

(a) Betweenness



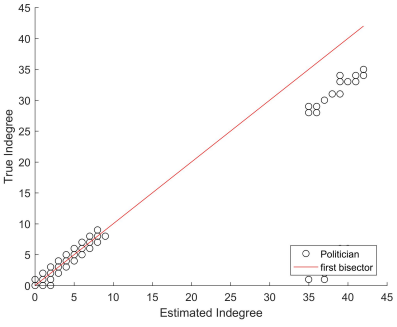

(b) Eigenvalue



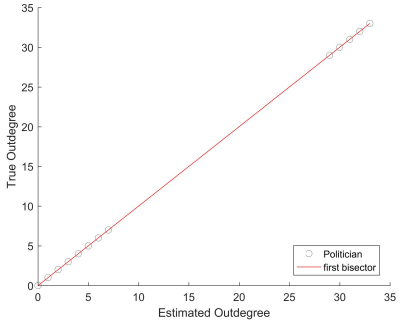

(c) Closeness




(d) Clustering



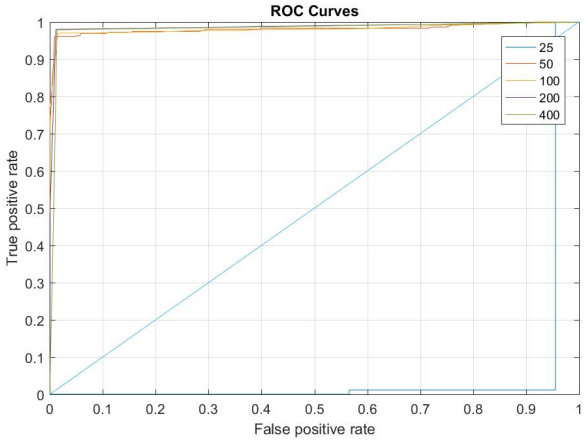

(e) Indegree




(f) Outdegree

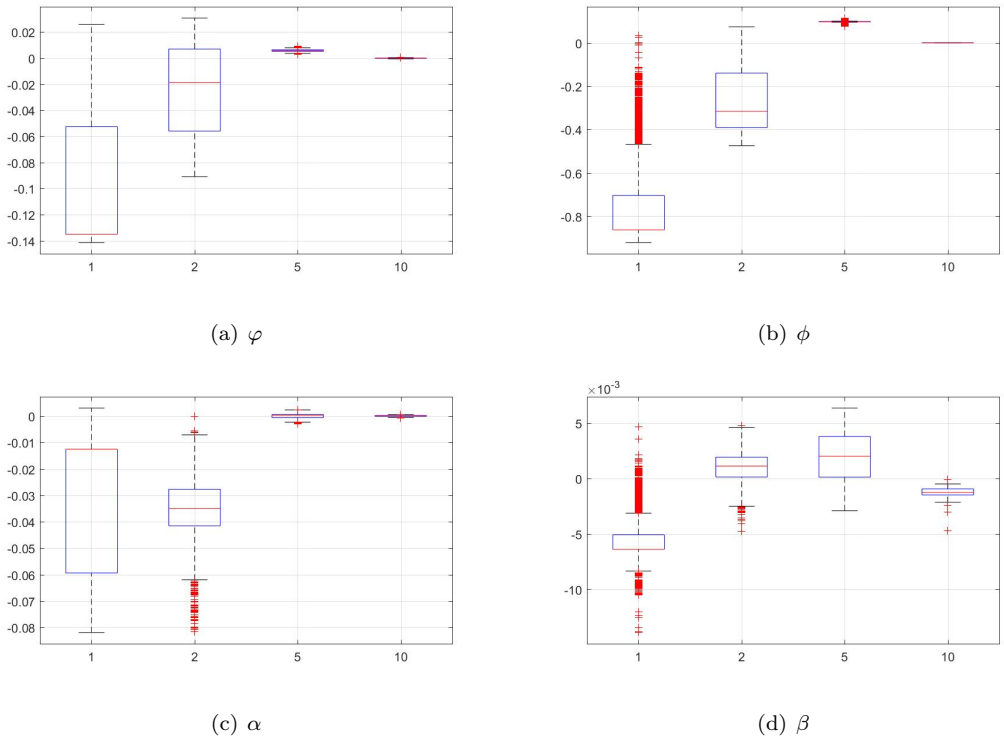
NOTE. X-axis: estimated value of node-level statistic. Y-axis: true value of node-level statistic. The true network is generated using equations (26)-(27) and data from the alumni network. The connections are generated from a circular network, as defined in Section 4.3.1. The DGP is described in detail in Section 8.5. The estimated network is derived using the parameter estimates at the last iteration of the MCMC. The first of $\bar{c} = 5$ networks is considered. See Newman [2010] for the definition of network measures.

Figure 7: GOODNESS OF FIT




NOTE. 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. Y-axis: the true positive rate (TPR) at various thresholds. X-axis: the false positive rate (FPR) at various thresholds. 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. The estimated network is derived using the parameter estimates at the last iteration of the MCMC. The first of $\bar{c} = 5$ networks is represented.

Figure 8: ESTIMATION BIAS FOR DIFFERENT NUMBERS OF REPEATED OBSERVATIONS OF VECTOR $E$ - PARAMETERS

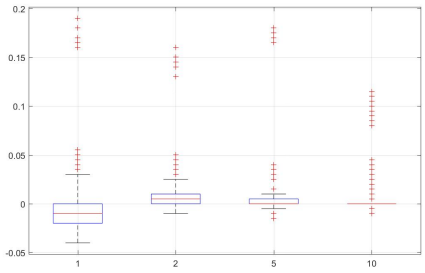



(a) $\varphi$ (b) $\phi$

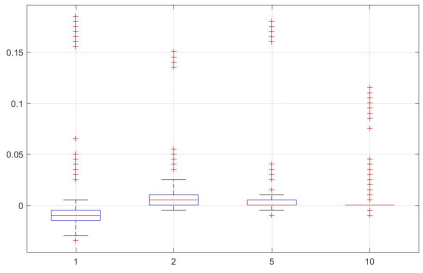
(c) $\alpha$ (d) $\beta$

NOTE. X-axis: number of repeated observations of vector $E$. Y-axis: distribution of the differences between estimated and true values in the MCMC after a burning period of 10,000 iterations. The estimated values are taken from the posterior distribution. The true values are fixed and generated using equations (26)-(27). The connections are generated from a circular network, as defined in Section 4.3.1. 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.

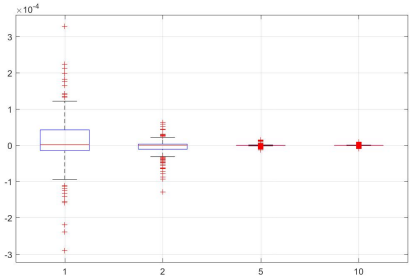
Figure 9: ESTIMATION BIAS FOR DIFFERENT NUMBERS OF REPEATED OBSERVATIONS OF VECTOR $E$ - NODE-LEVEL STATISTICS



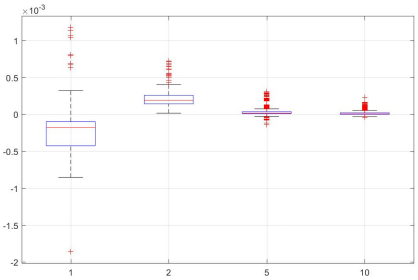

(a) Indegree



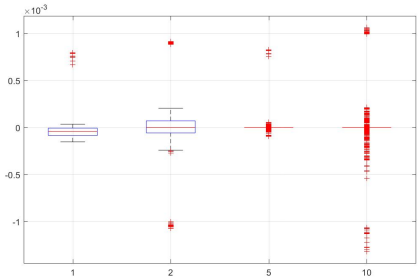

(b) Outdegree



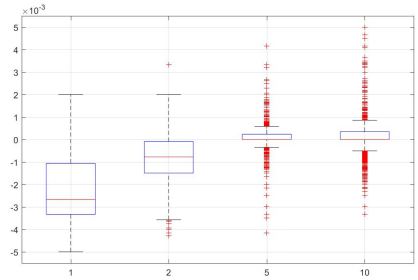

(c) Betweenness




(d) Closeness



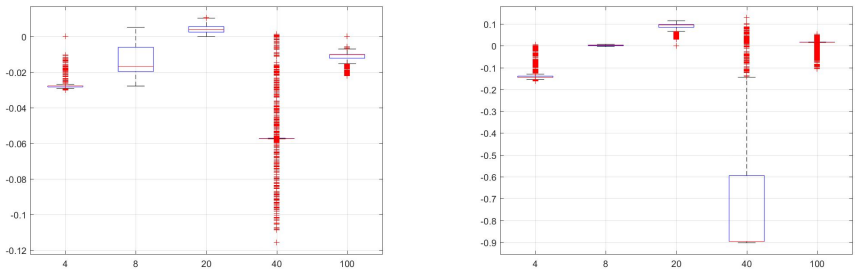

(e) Eigenvector




(f) Clustering

NOTE. X-axis: number of repeated observations of vector $E$. Y-axis: distribution of the differences between estimated and true values. See Newman [2010] for the definition of network measures. The true network is generated using equations (26)-(27). The connections are generated from a circular network, as defined in Section 4.3.1. The DGP is described in detail in Section 8.5. The estimated network is derived using the parameter estimates at the last iteration of the MCMC. Eigenvector centrality is not reported for the 25 nodes sample because the adjacency matrix has degenerate eigenvalues (i.e. with multiplicity greater than one). 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.

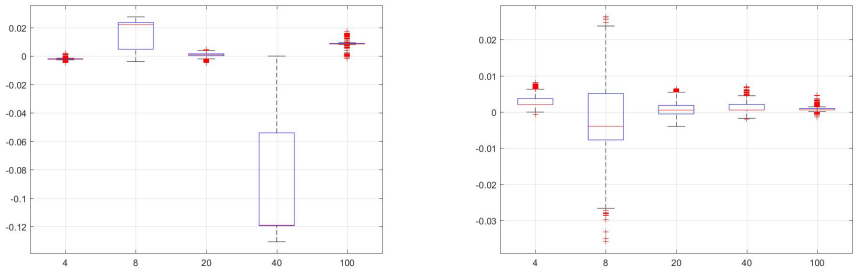
Figure 10: ESTIMATION BIAS FOR DIFFERENT NETWORK DENSITIES
- PARAMETERS -




(a) $\varphi$




(b) $\phi$



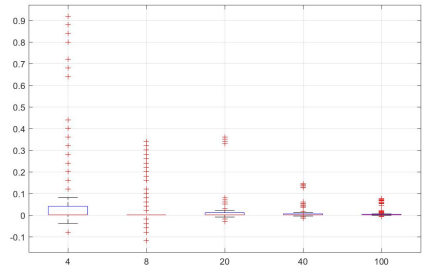

(c) $\alpha$



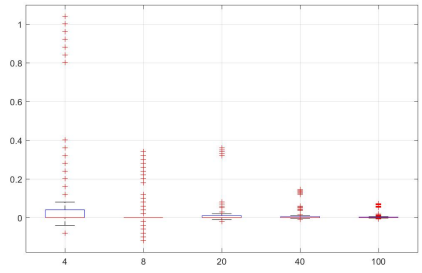

(d) $\beta$

NOTE. X-axis: network density is the maximum number of possible connections with other nodes ($\bar{z}$). Y-axis: distribution of the differences between estimated and true values in the MCMC after a burning period of 10,000 iterations. The estimated values are taken from the posterior distribution. The true values are fixed and generated using equations (26)-(27). The connections are generated from a circular network, as defined in Section 4.3.1. 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.

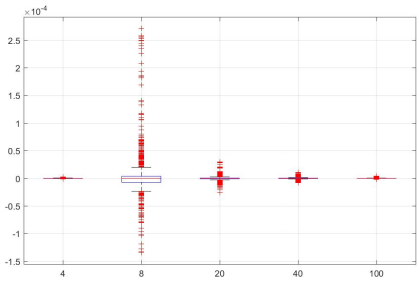
Figure 11: ESTIMATION BIAS FOR DIFFERENT NETWORK DENSITIES
- NODE-LEVEL STATISTICS -



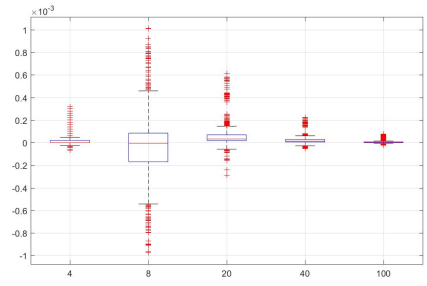

(a) Indegree



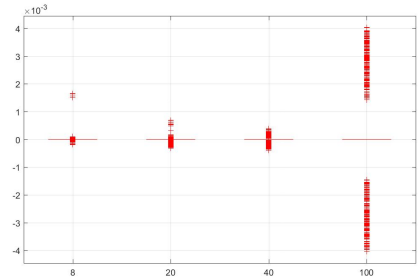

(b) Outdegree



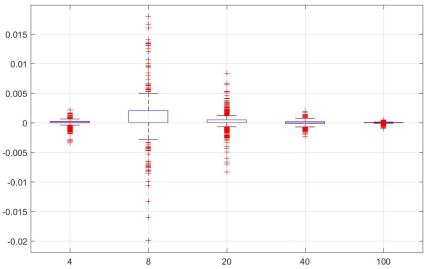

(c) Betweenness




(d) Closeness



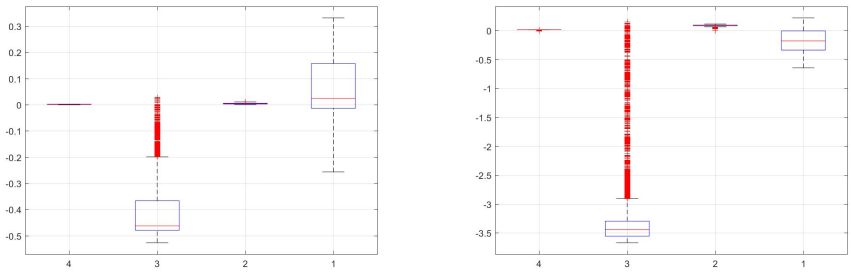

(e) Eigenvector




(f) Clustering

NOTE. X-axis: network density is the maximum number of possible connections with other nodes ($\bar{z}$). Y-axis: distribution of the differences between estimated and true values. See Newman [2010] for the definition of network measures. The true network is generated using equations (26)-(27). The connections are generated from a circular network, as defined in Section 4.3.1. The DGP is described in detail in Section 8.5. The estimated network is derived using the parameter estimates at the last iteration of the MCMC. Eigenvector centrality is not reported for the 25 nodes sample because the adjacency matrix has degenerate eigenvalues (i.e. with multiplicity greater than one). 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.

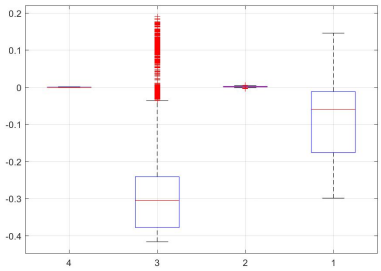
Figure 12: ESTIMATION BIAS FOR DIFFERENT $\phi$
- PARAMETERS -



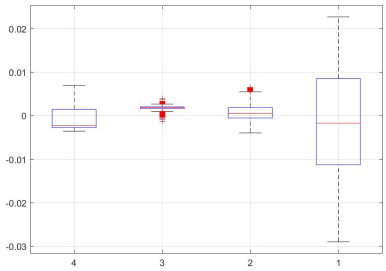

(a) $\varphi$




(b) $\phi$



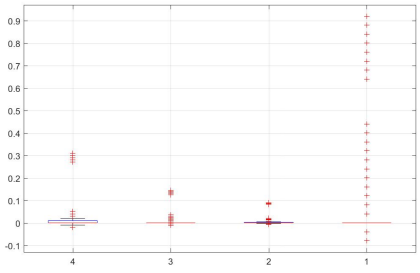

(c) $\alpha$



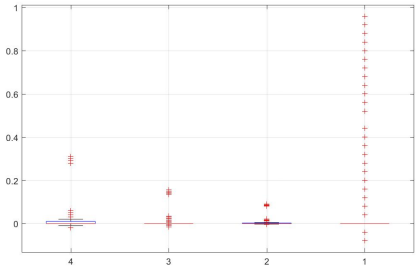

(d) $\beta$

NOTE. X-axis: $\phi$. Y-axis: distribution of the differences between estimated and true values in the MCMC after a burning period of 10,000 iterations. The estimated values are taken from the posterior distribution. The true values are fixed and generated using equations (26)-(27). The connections are generated from a circular network, as defined in Section 4.3.1. 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.

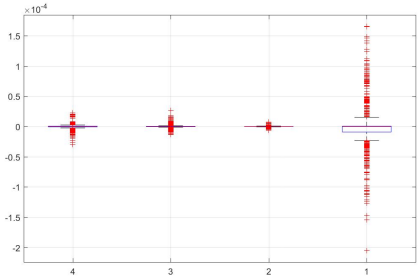
Figure 13: ESTIMATION BIAS FOR DIFFERENT $\phi$
- NODE-LEVEL STATISTICS -



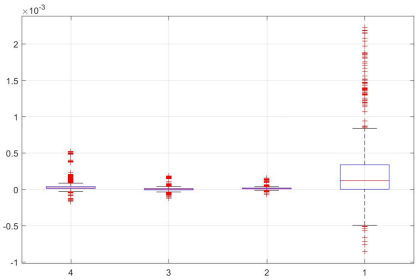
(a) Indegree



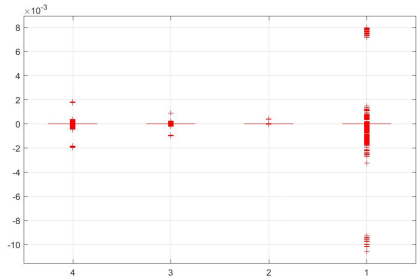
(b) Outdegree



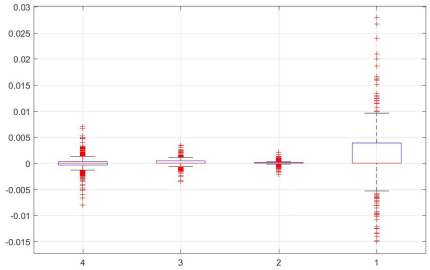
(c) Betweenness



(d) Closeness



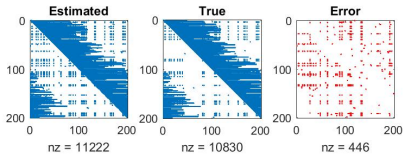
(e) Eigenvector



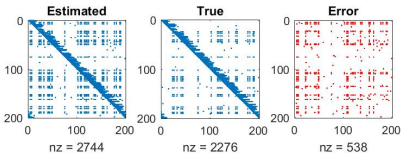
(f) Clustering

NOTE. X-axis: $\phi$. Y-axis: distribution of the differences between estimated and true values. See Newman [2010] for the definition of network measures. The true network is generated using equations (26)-(27). The connections are generated from a circular network, as defined in Section 4.3.1. The DGP is described in detail in Section 8.5. The estimated network is derived using the parameter estimates at the last iteration of the MCMC. Eigenvector centrality is not reported for the 25 nodes sample because the adjacency matrix has degenerate eigenvalues (i.e. with multiplicity greater than one). 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.

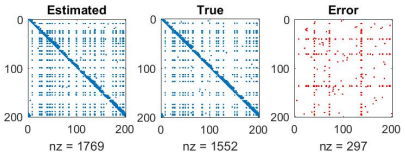
Figure 14: ESTIMATED VS TRUE NETWORKS - DIFFERENT TOPOLOGIES -



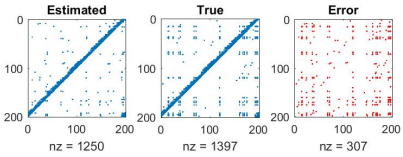

(a) High density ($\bar{z} = 100$)



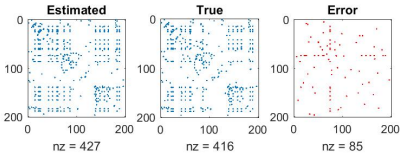

(b) Medium density ($\bar{z} = 20$)




(c) High $\phi = 4$



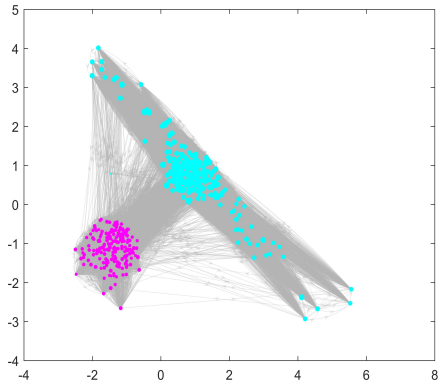

(d) Low density ($\bar{z} = 8$) and reverse circle



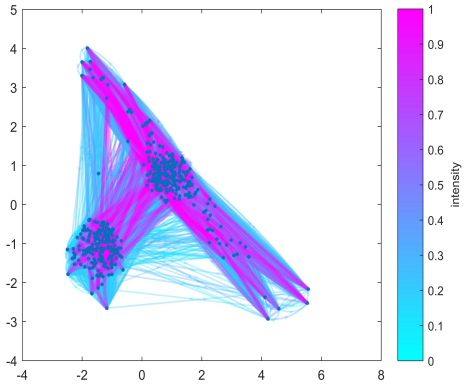

(e) Alumni network

NOTE. Adjacency matrices of true versus estimated networks (with blue dots) and their difference (with red dots). The true network is generated using equations (26)-(27) The DGP is described in detail in Section 8.5. The estimated network is derived using the parameter estimates at the last iteration of the MCMC. The first of $\bar{c} = 5$ networks is represented. The true networks in panels (a), (b), and (c) are generated with circular connections, as described in Section 4.3.1. The reverse linking network (in panel (d)) is obtained by reverting the algorithm described in Section 4.3.1, i.e. selecting links for i starting from i - 1 to i + 1 (i excluded) with a pace of -1. In the alumni network (panel (e)), the connections of 200 politicians from the 111th Congress extracted at random are considered. The 111th Congress is randomly selected.

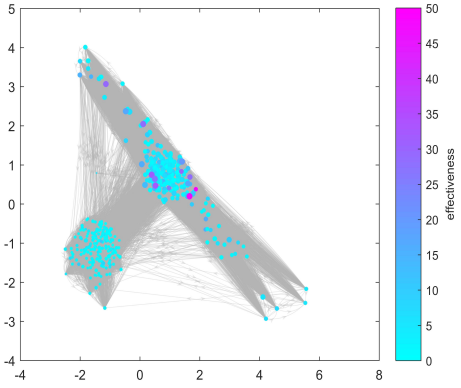
Figure 15: ESTIMATED NETWORK



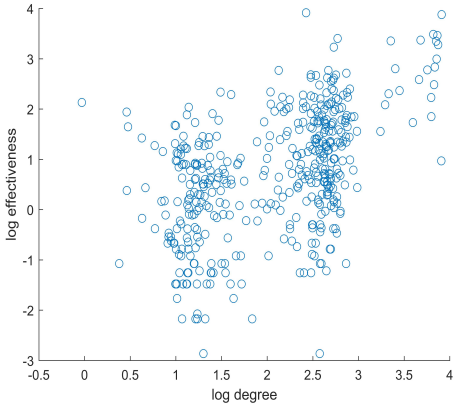
(a) Party




(b) Connections' strength




(c) Effectiveness




(d) Degree vs effectiveness

NOTE. The estimated network is derived using the parameter estimates at the last iteration of the MCMC for the 111th Congress. In panel (a) the color represents the party of the politician. Pink nodes are Republicans. In panel (b) the color of links is proportional to $g_{ij}$, the strenght of the connection between the two politicians. In panel (c) the color of each node is proportional to effectiveness, more pink nodes are more effective politicians. In panel (d) each point is a politician. The x-axis represents the (log) degree, the y-axis represents the (log) effectiveness. The estimated network is 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.

# 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.

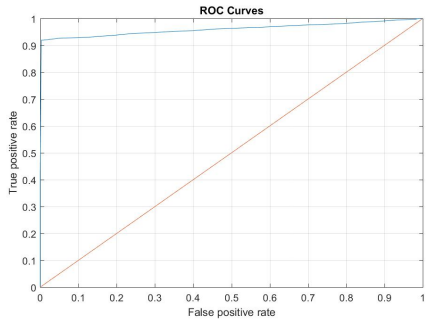
### A.1 Appendix Tables and Figures

Table A.1: SUMMARY STATISTICS

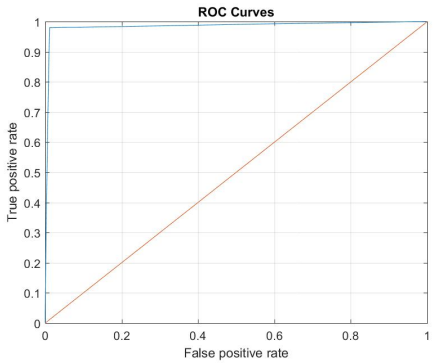
| Variable name | Variable definition | Mean | St. Dev |
|---|---|---|---|
| Party | Dummy variable taking value of one if the Congress member is a Democrat. | 0.5060 | 0.5019 |
| Gender | Dummy variable taking value of one if the Congress member is female. | 0.1723 | 0.3778 |
| Non white | Dummy variable taking value of one if the member of Congress is African-American or Hispanic, and zero otherwise. | 0.1388 | 0.3458 |
| Seniority | Maximum consecutive years in Congress. | 5.7863 | 4.4388 |
| Seniorty2 | Maximum 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 members state. | 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 | States level of professionalism [Squire, 1992]. | 0.1210 | 0.1779 |
| N. Obs. | | 2,176 | |

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

Figure A.1: NETWORK ESTIMATION - GOODNESS OF FIT




(a) Real network




(b) Circular network

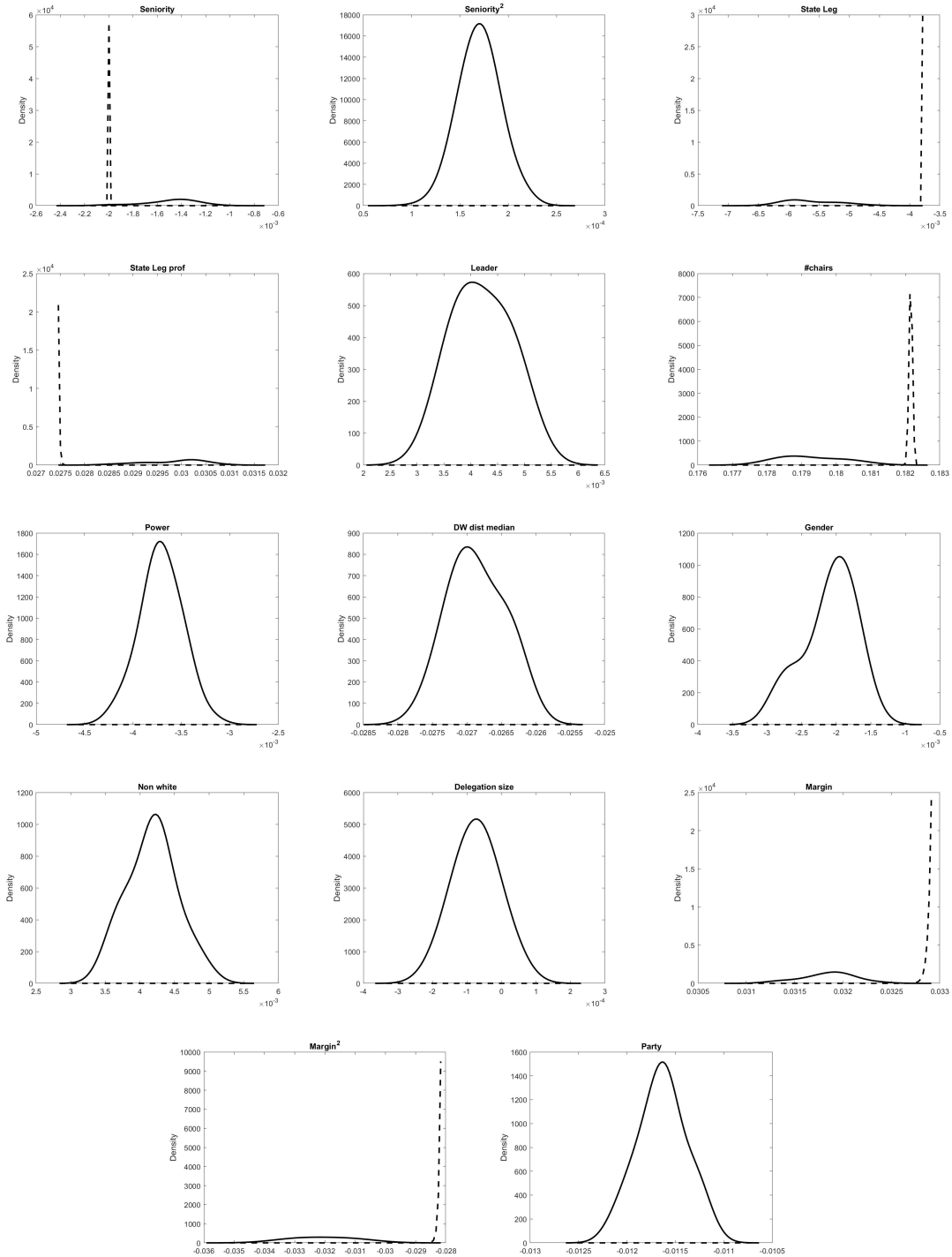
NOTE. Receiver Operating Characteristic (ROC) curve. The ROC curve is a plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied. Y-axis: the true positive rate (TPR) at various thresholds. X-axis: the false positive rate (FPR) at various thresholds. 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. The estimated network is derived using the parameter estimates at the last iteration of the MCMC. The first of $\bar{c} = 5$ networks is represented.

Table A.2: ESTIMATION RESULTS WITH UNOBSERVABLES

| Dependent variable: Legislative Effectiveness | | |
|---|---|---|
| | ABC<br>Endogenous network effects<br>(1) | ABC<br>Endogenous network effects<br>with unobservables<br>(2) |
| $\varphi$ | 0.0263 ***<br>[1.0000] | 0.1237 ***<br>[1.0000] |
| $\alpha$ | 0.7807 ***<br>[1.0000] | 0.4576 ***<br>[1.0000] |
| $\phi$ | 0.0741 ***<br>[1.0000] | 0.0231 ***<br>[1.0000] |
| Party | -0.0116 ***<br>[0.0000] | -0.0086 ***<br>[0.0000] |
| Gender | -0.0020 ***<br>[0.0000] | 0.0035 ***<br>[1.0000] |
| Non White | 0.0042 ***<br>[1.0000] | 0.0038 ***<br>[1.0000] |
| Seniority | -0.0015 ***<br>[0.0000] | -0.0012 ***<br>[0.0000] |
| Seniorty2 | 0.0002 ***<br>[1.0000] | 0.0002 ***<br>[1.0000] |
| DW ideology | -0.0268 ***<br>[0.0000] | -0.0293 ***<br>[0.0000] |
| Margin | 0.0319 ***<br>[1.0000] | 0.0402 ***<br>[1.0000] |
| Margin$^2$ | -0.0319 ***<br>[0.0000] | -0.0304 ***<br>[0.0000] |
| Committee Chair | 0.1789 ***<br>[1.0000] | 0.1809 ***<br>[1.0000] |
| Powerful Committee | -0.0037 ***<br>[0.0000] | -0.0053 ***<br>[0.0000] |
| Delegation size | -0.0001<br>[0.1240] | -0.0001 **<br>[0.0160] |
| Leader | 0.0040 ***<br>[1.0000] | 0.0111 ***<br>[1.0000] |
| State Legislative Experience | -0.0058 ***<br>[0.0000] | -0.0027 ***<br>[0.0000] |
| State Legislative Experience *<br>State Legislative Professionalism | 0.0300 ***<br>[1.0000] | 0.0304 ***<br>[1.0000] |
| $\mu_1$ | - | -0.0012 ***<br>[0.0000] |
| $\mu_2$ | - | 0.0002 ***<br>[1.0000] |
| $\mu_3$ | - | -0.0027 ***<br>[0.0000] |
| $\mu_4$ | - | 0.0304 ***<br>[1.0000] |
| $\mu_5$ | | 0.0111 ***<br>[1.0000] |
| Congress fixed effects | Yes | Yes |
| N. Obs. | 2,176 | 2,176 |

NOTE. Estimates of parameters in equation (11). 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.1. *, **, *** indicate statistical significance at the 10, 5 and 1 percent levels, based on empirical p-values. Each parameter $\mu$ corresponds to the relative power of $\epsilon$ as $\eta$ is generated with $\eta_{i,r} = \sum_{l=1}^{5} \mu_l \epsilon_{i,r}^l$.

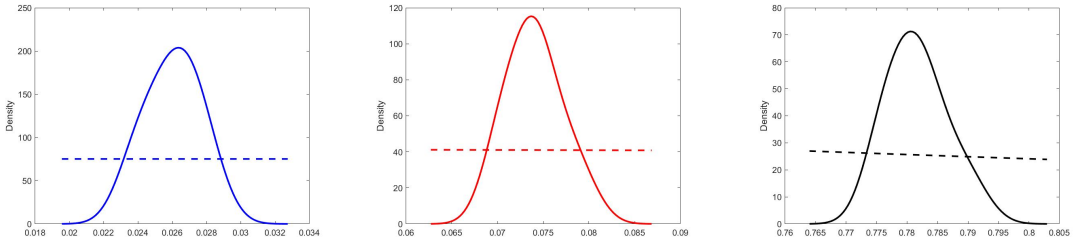
Figure A.2: ESTIMATED POSTERIOR DISTRIBUTIONS
- CONTROL VARIABLES -






NOTE. X-axis: parameter value, Y-axis: kernel density. The solid line represents the posterior, the dashed line the prior.

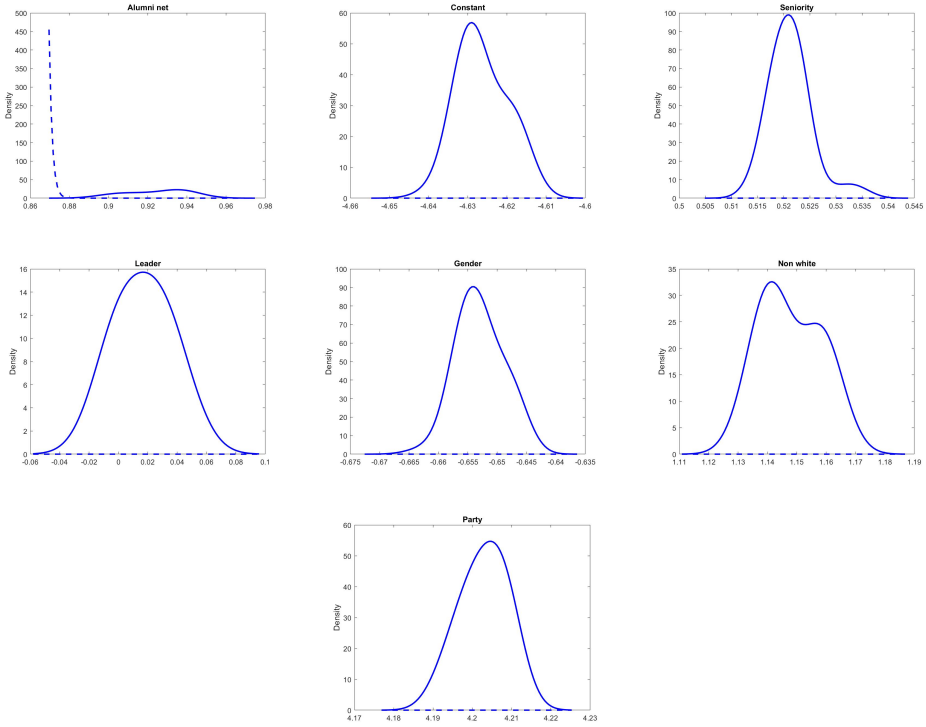
Figure A.3: ESTIMATED POSTERIOR DISTRIBUTIONS
- TARGET VARIABLES -




(a) $\varphi$ (b) $\phi$ (c) $\rho$

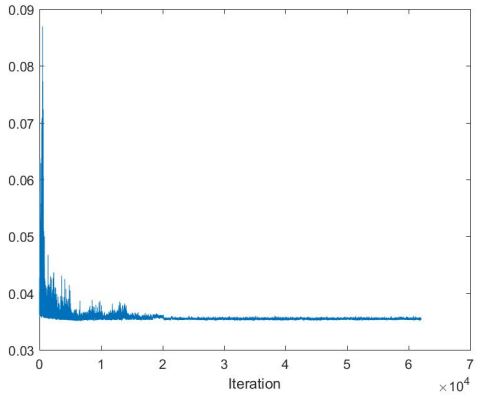
NOTE. X-axis: parameter value, Y-axis: kernel density. The solid line represents the posterior, the dashed line the prior.

Figure A.4: ESTIMATED POSTERIOR DISTRIBUTIONS
- LINK FORMATION -



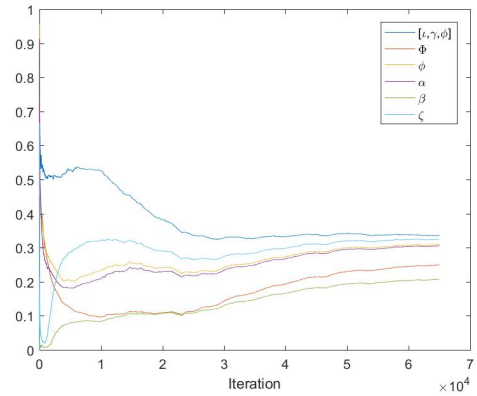

NOTE. X-axis: parameter value, Y-axis: kernel density. The solid line represents the posterior, the dashed line the prior.

Figure A.5: DISTANCE BETWEEN SIMULATED AND REAL DATA AT EACH ITERATION




NOTE. X-axis: MCMC iteration, Y-axis: distance value at each iteration.

Figure A.6: ACCEPTANCE RATE AT EACH ITERATION




NOTE. X-axis: MCMC iteration, Y-axis: frequency of acceptance: moving from $\omega$ to $\omega'$.

## A.2 A Comparison between the Estimated and Observed Networks

Here we present a comparison of the estimated network with other known and observable structures of interactions among politicians used in the literature to approximate political interconnectedness. Table A.3 reports some network-level statistics computed for the estimated network $G$ and the cosponsorship, the committee, and the alumni networks. It appears that the density in our estimated network is lower than the density of the cosponsorship and committee networks and higher than that of the alumni network, which are known for being too dense and too sparse, respectively.[57] The estimated network has closeness, clustering, and betweenness close to the cosponsorship network. The degree is more similar to the alumni network, while the assortativity is closer to the committee network. The level of assortativity, however, is significantly higher than in the other networks, suggesting a better ability to capture the fact that highly connected politicians are more likely to be connected among themselves. Clustering is also significantly higher than the other networks, plausibly better capturing the important role of political cliques. Betweenness and degree are significantly lower than the other networks, reflecting an estimated lower level of political intermediation and social connections.

To further analyze how different is the estimated network from the actual ones, we report the densities of degree, closeness, clustering, and eigenvector centralities in Figure A.7-A.9 for 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.4 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, while we cannot

[57] 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.

reject the hypothesis that the mean values of the centrality measures are the same in the estimated and actual networks in many cases, when looking at their distributions 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.

Table A.3: COMPARISON WITH OTHER NETWORKS
- NETWORK-LEVEL STATISTICS -

| | NETWORKS | | | |
|---|---|---|---|---|
| | ESTIMATED (1) | COSPONSORSHIP (2) | COMMITTEE (3) | ALUMNI (4) |
| Density | 0.0809 | 0.1069 | 0.2109 | 0.0036 |
| Assortativity | 66.4410 | 20.3642 | 47.4155 | 7.5099 |
| Diameter | Inf | Inf | Inf | Inf |
| Average distance | Inf | Inf | Inf | Inf |
| Closeness | 0.1403 | 0.1533 | 0.5768 | 0.0216 |
| Betweenness | 0.0001 | 0.0004 | 0.0694 | 0.0255 |
| Degree | 0.0304 | 0.0904 | 0.2262 | 0.0400 |
| Clustering | 0.7163 | 0.6561 | 0.6150 | 0.5410 |

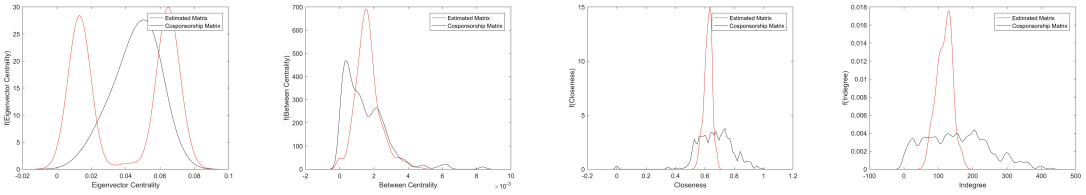
NOTE. The direct networks (cosponsorship and estimated) are transformed to indirect unweighted networks to have a clean comparison with the others. Given the direct network $D = \{d_{ij}\}$, its indirect unweighted counterpart is $U = \{u_{ij}\}$, where $u_{ij} = 1$ if $d_{ij}$ or $d_{ji}$ is different from zero, and zero otherwise. The network statistics are compared on a pooled network of five Congresses. See Newman [2010] for the definition of network statistics. The alumni network is defined in Section 4.1. See Battaglini and Patacchini [2019] for a description of the cosponsorship network. The $ij_{th}$ element of the committee network entry is equal to the number of Congressional committees in which both $i$ and $j$ sit.

Table A.4: NETWORK DIFFERENCES - STATISTICAL TESTS

| | Estimated Mean | Mean | T stat | p-value | Kolmogorov Smirnov | p-value |
|---|---|---|---|---|---|---|
| Cosponsorship | | | | | | |
| Indegree | 118.1162 | 153.8747 | -0.3919 | 0.6523 | 0.4465 | 0.0000 |
| Clustering | 0.7036 | 0.6816 | 0.2282 | 0.4098 | 0.2323 | 0.0000 |
| Betweenness | 0.0017 | 0.0015 | 0.1155 | 0.4540 | 0.2847 | 0.0000 |
| Closeness | 0.6220 | 0.6705 | -0.4032 | 0.6565 | 0.4784 | 0.0000 |
| Bonacich | 0.0401 | 0.0458 | -0.1827 | 0.5725 | 0.4305 | 0.0000 |
| Committee | | | | | | |
| Indegree | 172.4920 | 92.3554 | 1.9085 | 0.0285 | 0.8109 | 0.0000 |
| Clustering | 0.7036 | 0.7115 | -0.0376 | 0.5150 | 0.3781 | 0.0000 |
| Betweenness | 0.0017 | 0.0021 | -0.1002 | 0.5399 | 0.3622 | 0.0000 |
| Closeness | 0.6969 | 0.5768 | 1.5566 | 0.0601 | 0.8360 | 0.0000 |
| Bonacich | 0.0401 | 0.0393 | 0.0226 | 0.4910 | 0.2642 | 0.0000 |
| Alumni | | | | | | |
| Indegree | 172.4920 | 1.5626 | 6.2572 | 0.0000 | 1.0000 | 0.0000 |
| Clustering | 0.7036 | 0.1575 | 1.6721 | 0.0476 | 0.8588 | 0.0000 |
| Betweenness | 0.0017 | 0.0009 | 0.2264 | 0.4105 | 0.8018 | 0.0000 |
| Closeness | 0.6969 | 0.0216 | 14.3031 | 0.0000 | 1.0000 | 0.0000 |
| Bonacich | 0.0401 | 0.0109 | 0.5195 | 0.3018 | 0.9089 | 0.0000 |

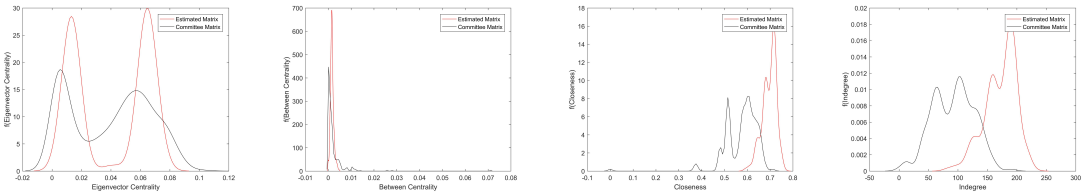
NOTE. Node-level statistics are considered. See Newman [2010] for the definition of network statistics. The first four columns test differences in means, the last two columns test the difference between the two distributions. The alumni network is defined in Section 4.1. Cosponsorship activity is measured 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 entry 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.

Figure A.7: ESTIMATED VS COSPONSORSHIP NETWORK
- DENSITIES OF NODE-LEVEL STATISTICS -



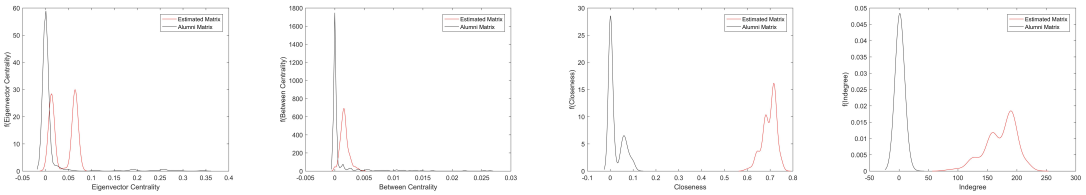

(a) Eigenvector (b) Closeness (c) Clustering (d) Degree

Figure A.8: ESTIMATED VS COMMITTEE NETWORKS
- DENSITIES OF NODE-LEVEL STATISTICS -




(a) Eigenvector (b) Closeness (c) Clustering (d) Degree

Figure A.9: ESTIMATED VS ALUMNI NETWORKS
- DENSITIES OF NODE-LEVEL STATISTICS -




(a) Eigenvector (b) Closeness (c) Clustering (d) Degree

NOTE. 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 each measure. The estimated network is derived using the parameter estimates at the last iteration of the MCMC for the 111th Congress. The alumni network is defined in Section 4.1. The $ij_{th}$ element of the committee network entry 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.