AN EQUILIBRIUM MODEL OF INTER-ORGANIZATIONAL NETWORK FORMATION

SHWETA GAONKAR AND ANGELO MELE

Abstract. This paper studies the formation of inter-organizational ties in an equilibrium framework. We model firms' decisions to form links with other firms as a strategic game. Firms' payoffs include costs and benefits of establishing a relationship among firms, as well as equilibrium effects from transitivity. We characterize the equilibrium networks as exponential random graphs (ERGM) and we estimate the firms' payoffs using a Bayesian approach. We apply the framework to a co-investment network of venture capital firms in the medical device industry. The results show how controlling for the endogenous network structure in equilibrium is crucial to provide a good fit of real-world network data. Firms' payoffs show a preference for links to similar firms (homophily) and that have common partners (transitivity). Our structural approach allows us to model the effects of economic shocks or changes on the network. We show how entry of new firms or minimum capital requirements increase the density and clustering of the co-investment network, thus allowing us to quantify the equilibrium impact of these market shocks.

Keywords: Structural model, tie formation, exponential random graphs, homophily, clustering, bayesian estimation, and VC syndication

This version: December 20, 2018. First version: December 15, 2017. Contact: gaonkar@jhu.edu and angelo.mele@jhu.edu. We thank Maqbool Dada, Corey Phelps for comments and suggestions.

1. Introduction

Inter-organizational ties are crucial for a firms' success (Lavie (2007)) as they provide access to resources required to deal with competition, and create value (Gulati (1995b), Powell et al. (1996)). In order to disentangle the relationship between networks, and firm outcomes (Ahuja et al. (2009), Stern et al. (2014), Martin et al. (2015)), empirical researchers have extensively examined the formation of these ties (Ahuja (2000), Ahuja et al. (2009), Chung et al. (2000), Garcia-Pont and Nohria (2002), Gulati (1999),Kogut et al. (2007), Li and Rowley (2002), Rothaermel and Boeker (2008), Hallen (2008), Reuer and Lahiri (2013), Stern et al. (2014)). However, a key challenge of studying the formation of ties is to address the interdependence among these relationships.

This study addresses this challenge by proposing a generalized structural model that captures the network formation process among firms. A structural model provides a rigorous approach to understand the network formation process. More importantly, it allows the researchers to simulate counter-factual policy experiments that predict the changes in network structure in response to any change in the environment surrounding the firm. The ability to predict the emergence of network structure allows management scholar to deduce important managerial implications based on strategic decisions made by firms, or change in regulation.

We demonstrate the use of the structural model by applying this approach to the network of co-investments of venture capital firms in the medical devices industry, using data from 1940 to 2013. Two venture capital companies are considered linked if they have co-invested in the same medical devices start-up. The network exhibits a standard core-periphery structure, where a few firms have many connections among themselves (the core), while the rest of the firms have few links to the main core (the periphery). The structural model allows us to understand what factors lead to this network structure. The estimated results strongly support the homophily argument established by the extant literature on VC syndication networks (Gulati and Gargiulo (1999), Kogut et al. (2007), Zhang and Hallen (2017)). The results suggest that firms tend to form ties with firms that are similar to them in age, geographic location, and managed capital.

An additional advantage of our approach is that we can provide tests for goodness of fit of our model. To verify our results, we check model fit for our main specification that includes endogenous network terms, which we compare to the fit of a model without endogenous network terms. The latter model corresponds to a logit, where links are assumed independent. The model with endogenous network terms can replicate most features of the observed data, including degree distribution and geodesic distance distribution. On the other hand, the model without endogenous network terms fails to match most of these features. We conclude that our equilibrium modeling of network formation is crucial to replicate real-world features of the data and provides an economic interpretation of the parameters. This feature of our model allows us to understand firms' incentives to form ties while controlling for endogeneity generated by the strategic equilibrium.

More importantly, our structural equilibrium approach allows us to run counterfactual experiments. Our model explicitly accounts for the equilibrium network effects on individual firms decisions to form and cut ties. When there is a policy change or an exogenous shock to the economy, firms will make an optimal decision regarding their new partnerships and alliances, thus changing the shape of their network. However, firms will also respond optimally to the decisions of other firms, as a consequence of the policy change or exogenous shock, triggering an additional secondary adjustment of their strategy following these equilibrium effects. Our equilibrium model takes this into account, while other partial equilibrium models are not able to incorporate the equilibrium adjustments. We focus on two types of policy changes; 1) entry of new firms in the market; 2) imposition of a minimum capital to operate in these markets. We show how these policies change the equilibrium network configuration. In general, we observe that networks, in response to these shocks become more clustered and dense. Entry of new firms or minimum capital requirement affect the degree distribution: many firms that did not have any links before the entry or regulatory change will form at least one alliance, thus making the network denser and contributing to the increase in clustering. The new links are not randomly allocated, as we document an average increase in homophily: thus firms tend to form more links to similar firms after the policy changes.

2. Methods for Predicting Formation of Inter-organizational Ties

Empirical works that have examined network formation have explored these relationships as dyadic or triadic level outcomes. Most dyadic- or triadic-level studies examining tie formation have used binary choice models to analyze these relationships (Shipilov and Li (2012)), wherein firms select partners based on observable attributes, such as their resources, specialization (Chung et al. (2000), Rothaermel and Boeker (2008)), trust (Gulati (1995*a*)) or their network positions (Ahuja et al. (2009), Gulati and Gargiulo (1999)). However, existing research on embeddedness has shown that network ties are correlated (Gulati (1995*a*)), making each link formation decision endogenous. This stream of research has focused on how the firm's position in a preexisting network structure determines the formation of inter-organizational relationships (Gulati (1995*a*), Powell et al. (1996)). These preexisting ties create path dependence in the way firms establish new relationships with other firms because repeated interaction reduces uncertainty while increasing interdependence (Gulati and Gargiulo (1999)). In other words, when firms face competition and uncertainty, they rely on familiar partners they can trust (Beckman et al. (2004), Sorenson and Stuart (2008)).

The challenge of estimating network formation using binary choice models is that they fail to account for interdependence among ties. For example, when a firm forms several links, a discrete choice model like logit or probit would consider those decisions to be independent.[1] But, the

[1] For more details regarding the limitation and current methodology used in management literature, please refer to Kim et al. (2015) and Mindruta et al. (2016).

formation of several links consists of interrelated decisions if one assumes the company behaves strategically and each link is costly. As a consequence, the links should not be considered independent in statistical analysis. This view is supported by the embeddedness literature, which has established that prior ties strongly influence the choice of future partners (Gulati (1995*b*)), showing the importance of the underlying network structure as well as the observable actors' attributes in the formation of ties.

Some researchers in management who study firm networks have adopted new modeling and estimation techniques to overcome these challenges. One such approach involves the use of matching models (Mindruta et al. (2016), Fox et al. (forthcoming)). However, matching models are better suited for the analysis of bipartite networks, where the market can be divided into two sides. When we are evaluating a market in which each firm can form links to any other firm, we need a different modeling strategy. Few papers have examined the network formation process using ERGM. For example, Kim et al. (2015) studies board interlocks formation, using ERGMs and a frequentist approach, and estimate the model with the MCMC-MLE algorithm (Snijders (2002)).

Despite its advantage, recent work in economics and statistics has shown that some of ERGM features may limit their use in empirical applications of interest to economics and management. First, Bhamidi et al. (2011), Diaconis and Chatterjee (2013) and several practitioners have shown that many empirical specifications of ERGMs suffer from *degeneracy*; the simulation of such models generates degenerate networks, either almost empty or almost complete. This feature of the model makes it difficult to fit the observed data and makes estimators unstable and unreliable. Second, ERGMs are statistical models whose parameters may not have direct economic interpretation (Mele (2017*a*)). Third, the available estimation methods for ERGMs have some limitations. The Maximum Pseudo-Likelihood Estimator (MPLE) approximates the likelihood by assuming independence of the conditional link probabilities. While this estimator is consistent, practitioners have documented its poor performance in applied work using small- and medium-size networks (Snijders (2002), Caimo and Friel (2011)). On the other hand, the Markov Chain Monte Carlo Maximum Likelihood Estimator (MCMC-MLE) approximates the likelihood function by simulation, finding an approximate estimate. The accuracy of the approximation depends on meticulously choosing a simulation starting value that is close to the exact maximum likelihood estimate. Otherwise, the likelihood approximation is extremely noisy, and the optimization becomes unstable and unreliable (Chandrasekhar and Jackson (2014), Snijders (2002), Geyer and Thompson (1992)). Furthermore, in this class of models, depending on the data, the maximum likelihood estimate may not exist (Geyer and Thompson (1992)), leading to unreliable inference.

3. An equilibrium model of network formation

3.1. **Setup.** The economy consists of n firms and time is discrete ($t = 1, 2, 3, ...$). A generic firm i has a set of observable characteristics that we denote x_i. For example, x_i could include the size of the firm, its industry and location. We assume that each firm has M observable attributes, that is $x_i = \{x_{i,1}, x_{i,2}, ..., x_{i,M}\}$, and we denote as x the $n \times M$ matrix that contains all the vectors of firms' observable attributes.

The network of firms is represented by an $n \times n$ adjacency matrix g. The element at row i and column j will be denoted g_{ij}; we follow the convention in the literature and set $g_{ij} = 1$ if there is a link between firms i and j; otherwise $g_{ij} = 0$. In our application the network is *undirected* and each link requires mutual consent; therefore the adjacency matrix g is symmetric (that is, $g_{ij} = g_{ji}$ for all $i, j = 1, ..., n$). Most of the theoretical results below can be easily extended to directed networks.[2]

[2]See Mele (2017*a*) for a treatment of directed networks.

3.2. **The network formation game.** We assume that the network is formed sequentially over time and that firms maximize the surplus generated by each link. We distinguish between *opportunity* and *willingness* to form a link.

We model opportunity through a stochastic meeting process. In each period, two randomly selected firms, i and j, have a chance to form (or delete) a link. This opportunity occurs with probability $\rho(g, x_i, x_j)$, which can depend on the existing network g. For example, firms with shared partners may have more frequent chances to form alliances. Furthermore, the probability ρ can depend on the observable firms' characteristics x_i and x_j; for instance, firms with similar observable characteristics may have more opportunities to form partnerships.[3]

Upon receiving the opportunity to modify a link, firms i and j decide whether they want to update their connection g_{ij}. If the link does not exist, they decide whether to form a new link; if the link already exists, they choose whether to delete the link. When choosing whether to update the link, companies behave myopically, maximizing the current surplus generated by their link.[4]

To characterize the equilibrium of the model and obtain a tractable and estimable likelihood, we make some assumptions on the payoffs and the probabilities $\rho(g, x_i, x_j)$ that govern the rate at which firms receive opportunities to create and delete links. Let g_{-ij} denote the network g with the exclusion of link g_{ij}.

ASSUMPTION 1. *The meeting process is i.i.d. over time, the probability that firms i and j meet is*

$$\rho(g, x_i, x_j) = \rho(g_{-ij}, x_i, x_j) > 0 \tag{1}$$

and the sum of these probabilities over all possible pairs of firms is one.

[3]See Currarini et al. (2009) and Currarini et al. (2010) for a model where meetings are biased in favor of agents of the same group. Mele (2017*a*), Badev (2013) and Chandrasekhar and Jackson (2014) also consider variants of this "meeting" technology.

[4]This modeling approach has been used in previous work by Nakajima (2007), Mele (2017*a*), Mele and Zhu (2017), Badev (2013), Bala and Goyal (2000) and Jackson and Watts (2001) among others.

Assumption 1 guarantees that any pair of firms has an opportunity to update their links. The probability $\rho(g_{-ij}, x_i, x_j)$ can be very small, but it is necessarily positive. The main implication is that the network formation process can reach any equilibrium network with positive probability (Mele (2017*a*)). The simplest probability model for ρ that satisfy Assumption 1 is a discrete uniform distribution. More generally, the probability ρ can depend on the network g_{-ij}. The crucial part of Assumption 1 is that the probability for any pair is positive.

Firms' payoffs are defined over networks g and observable characteristics x. We will denote as $U_i(g, x; \theta)$ the payoff of firm i from network g, observable characteristics x, and parameters $\theta = \{\alpha, \beta, \gamma\}$.

ASSUMPTION 2. *The payoff of firm i is*

$$U_i(g, x; \theta) = \sum_{j=1}^{n} g_{ij} \left[u(x_i, x_j; \alpha) + \beta \sum_{r \neq i,j}^{n} g_{jr} + \gamma \sum_{r \neq i,j}^{n} g_{jr} g_{ri} \right] \tag{2}$$

where $u(x_i, x_j; \alpha)$ is a symmetric function.

The payoff $U_i(g, x; \theta)$ has three components. First, when firm i forms a link with firm j, it receives a net payoff $u(x_i, x_j; \alpha)$ that depends on characteristics x_i and x_j, and a parameter α. For example, a firm may find companies in the same industry more attractive for a partnership. The payoff $u(x_i, x_j; \alpha)$ includes both *costs and benefits* of direct connections, so it should be interpreted as *net direct benefit* of forming a link. We follow the convention in the strategic network literature and assume that *firms pay a cost for direct links, but that indirect connections are free* (Jackson and Wolinsky (1996), Jackson (2008)). In the empirical section we will be more explicit about the functional form of $u(x_i, x_j; \alpha)$.

Second, when firm i connects to a firm j, it receives an additional payoff β for each firm that has formed a link to j in previous periods. If firm j has 2 links, then firm i will receive 2β; if firm j has 5 links, then firm i will receive 5β. If $\beta > 0$, firm i prefers to link to a "popular" firm; vice-versa, there could be a competition effect, and i may receive less utility from a "popular" firm since this firm has to share resources with other partners. The sign and magnitude of β is ultimately an empirical question. Third, firm i receives a payoff of γ for each partner in common with j. The term $\sum_{r\neq i,j}^{n} g_{jr}g_{ri}$ corresponds to the number of common partners between i and j.

This part of the payoff captures the clustering effects or triadic closure process. When γ is positive, a firm receives more surplus from companies with which it shares many partners. Viceversa, when γ is negative a firm receives negative surplus when linking to a company that has many shared partners. In the network literature there is an empirical regularity: if two nodes have a common neighbor, there is a high chance that they form a link (Wasserman and Faust (1994), Wasserman and Pattison (1996), Jackson (2008)). This model accounts for this property through the payoff structure, generating the triadic closure property as an equilibrium outcome of the network formation game. On the other hand, we do not impose a positive parameter γ and instead let the data determine the sign. Therefore our model can accommodate both transitivity and intransitivity.

Finally, we assume that firms receive a joint matching shock $\varepsilon_{ij} = (\varepsilon_{0,ij}, \varepsilon_{1,ij})$ before choosing whether to update a link. The random shock models idiosyncratic reasons that could affect the decision to link. For example, in some periods two firms may be a bad match (negative matching shock) for reasons that are unobservable to the researcher, such as misaligned long-term strategies or incompatible risk profiles. On the other hand, there are periods in which those companies may be a good match (positive matching shock), increasing the willingness of both firms to create a partnership.

ASSUMPTION 3. *Firms receive a logistic matching shock before updating their links, which is i.i.d. over time and across pairs.*

Assumption 3 is standard in discrete choice models and random utility models in the empirical literature. In our model, this assumption is crucial to derive the likelihood of the network in closed-form,[5] allowing us to perform maximum likelihood estimation or Bayesian estimation.[6]

3.3. **Characterization of Equilibrium Networks.** We focus on equilibrium networks that satisfy *pairwise stability with transfers*, one of the most common equilibrium notions used in the network literature in economics.[7] This equilibrium notion requires that both firms consent to form a new link. However, the surplus generated by a new link is allowed to differ between the partners, because firms can transfer part of the payoff to the other party. This is a way to model asymmetric bargaining power or different investments of resources in the partnership.

We will denote the transfer from firm i to firm j as τ_{ij}. By definition, $\tau_{ij} = -\tau_{ji}$. Conditional on being randomly selected, firms i and j will form (or keep) a link if the surplus generated by forming the link (including transfers and matching shocks) is larger than the surplus without the links, that is if

$$\begin{aligned} &U_i(g_{ij}=1, g_{-ij}, x_i, x_j; \theta) + U_j(g_{ij}=1, g_{-ij}, x_j, x_i; \theta) + \varepsilon_{1,ij} \geq \\ &U_i(g_{ij}=0, g_{-ij}, x_i, x_j; \theta) + U_j(g_{ij}=0, g_{-ij}, x_j, x_i; \theta) + \varepsilon_{0,ij} \end{aligned} \tag{3}$$

In equation (3) the transfers do not appear; if i tranfers a positive amount τ_{ij} to j, then i's payoff will be $U_i(g_{ij}=1, g_{-ij}, x_i, x_j; \theta) - \tau_{ij}$ while for j the payoff will be $U_j(g_{ij}=1, g_{-ij}, x_j, x_i; \theta) + \tau_{ij}$. When we sum the payoffs to compute the total surplus of the link, the transfers cancel out. Therefore,

[5]See also Mele (2017*a*), Mele and Zhu (2017), Chandrasekhar and Jackson (2014) and Heckman (1978).

[6]As an alternative to assumptions 1-3, we could use the spatial GMM (Conley (1999), Conley and Topa (2007)) or the Approximate Bayesian Computation (ABC). These techniques do not require knowing the likelihood in closed-form (Marjoram et al. (2003), König (2016)).

[7]See Jackson (2008), Mele and Zhu (2017) and Chandrasekhar and Jackson (2014) for examples.

the condition (3) can be re-stated as follows

$$U_i(g_{ij}=1,g_{-ij},x_i,x_j;\theta)+U_j(g_{ij}=1,g_{-ij},x_j,x_i;\theta)-$$
$$[U_i(g_{ij}=0,g_{-ij},x_i,x_j;\theta)+U_j(g_{ij}=0,g_{-ij},x_j,x_i;\theta)]\geq$$
$$\varepsilon_{0,ij}-\varepsilon_{1,ij}$$

Let's define

$$\Delta_{ij}:=U_i(g_{ij}=1,g_{-ij},x_i,x_j;\theta)+U_j(g_{ij}=1,g_{-ij},x_j,x_i;\theta)-$$
$$[U_i(g_{ij}=0,g_{-ij},x_i,x_j;\theta)+U_j(g_{ij}=0,g_{-ij},x_j,x_i;\theta)]$$

Our Assumption 3 implies that the matching shocks are logistic. Therefore the difference of matching shocks $\varepsilon_{0,ij}-\varepsilon_{1,ij}$ is also logistic. The main consequence is that the probability that firms i and j form (or keep) a link in any period has the standard logit form (Heckman (1978))

$$P(g_{ij}=1|g_{-ij},x_i,x_j,\theta)=\frac{\exp[\Delta_{ij}]}{1+\exp[\Delta_{ij}]} \tag{4}$$

Equation (4) is a *conditional* probability. It describes the probability of a link between i and j given the rest of the links in the network g_{-ij}, the observable characteristics x_i and x_j and the parameters of the model θ. This probability is also implicitly conditioned on the event that firms i and j have an opportunity to revise their link, which happens with probability $\rho(g_{-ij},x_i,x_j)$. The probability (4) shows how the link g_{ij} depends on the other links in the network g_{-ij}, thus relaxing the independence assumption implicit in a standard logit model. Indeed, this is what strategic network formation implies: there is dependence among linking decisions, induced by the strategic equilibrium.

The model generates a sequence of networks as a result of link creation or deletion. This sequence is Markovian and converges to a unique stationary distribution that characterizes the probability of observing a specific network architecture in the long run, as shown in Mele (2017*a*).

Let's define the aggregate function

$$Q(g,x;\theta)=\sum_{i=1}^{n}\sum_{j=1}^{n}g_{ij}u(x_i,x_j;\alpha)+\frac{\beta}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{r\neq i,j}^{n}g_{ij}g_{jr}+\frac{\gamma}{6}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{r\neq i,j}^{n}g_{ij}g_{jr}g_{ri} \tag{5}$$

Equation (5) is called a *potential function*, and it summarizes the incentives of each firm to form links, net of the matching shock. The crucial property of the potential is

$$Q(g,x;\theta)-Q(g',x;\theta)=U_i(g,x;\theta)+U_j(g,x;\theta)-\left[U_i(g',x;\theta)+U_j(g',x;\theta)\right] \tag{6}$$

where g is a network where firms i and j have a link (that is $g_{ij}=1$), and g' is the same network g, excluding the link between firms i and j (that is $g'_{ij}=0$ and $g'_{-ij}=g_{-ij}$). Notice that the right-hand side of (6) represents the incentive of firms i and j to form the link; if the sum of their payoffs when they form the link is greater than the sum of their payoffs when they do not have a link, then they will form the link (excluding the stochastic matching shock). The left-hand side of (6) shows that this difference in payoffs can be retrieved using the potential function. This property holds for any pair of firms i and j and for any link g_{ij}, $i,j=1,...,n$.[8]

This means that the profitable deviations of firms i and j can be computed using the difference in potential functions (the left-hand side of equation (6)). Therefore all the equilibrium networks can be found using the potential.

A network is *pairwise stable with transfers* if no two firms want to form a link or delete a link. In our model, all the *pairwise stable networks (with transfers)* correspond to the (local) maxima of the potential function (5).[9] Intuitively, let's consider a network that maximizes the potential function. Suppose we delete the link between firms i and j. This will decrease the potential because the network was a maximizer of the potential. A decrease in potential means that the sum of payoffs of i and j is higher when the link exists than it is when we delete their link. Therefore, they will not be willing to delete their partnership. The same will hold if we consider an additional

[8]The formal proof is contained in Mele (2017*a*) and Mele and Zhu (2017).

[9]See Monderer and Shapley (1996), Mele (2017*a*), Jackson and Watts (2001) and Badev (2013).

link between two firms in the network. Adding a link will decrease the potential and therefore implies that the firms involved in this relationship are better off not forming the new link. We can repeat this reasoning for any pair of firms, showing that indeed the network that maximizes the potential function is a pairwise equilibrium network with transfers. This is an important result that facilitates the computation of the equilibria. Furthermore, the existence of at least a pairwise stable equilibrium with transfers is guaranteed by the existence of the potential function, as shown in Monderer and Shapley (1996).

We should note that the potential function is different from a welfare function. Indeed if we assume a classical welfare function, which is the sum of the firms' payoffs, we have

$$
\begin{aligned}
W(g,x;\theta) &= \sum_{i=1}^{n} U_i(g,x;\theta) && (7)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}u(x_i,x_j;\alpha) + \beta\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{r\neq i,j}^{n} g_{ij}g_{jr} + \gamma\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{r\neq i,j}^{n} g_{ij}g_{jr}g_{ri} && (8)
\end{aligned}
$$

As a consequence, the pairwise stable networks are not necessarily efficient. This is a classical result in the economics of networks literature. The decentralized equilibrium is not efficient because there are externalities in link formation; a firm that forms a direct link pays a cost and receives some benefits, while an indirect link has no cost but brings additional benefits (positive or negative). Therefore, the formation of a new link creates an externality for other firms that could benefit from a new indirect connection without paying any cost. This externality can be positive or negative, depending on the sign of the coefficients β and γ.

Using the potential function characterization we can write the conditional probability of linking as

$$P(g_{ij}=1|g_{-ij},x_i,x_j,\theta) = \frac{\exp\left[\Delta_{ij}\right]}{1+\exp\left[\Delta_{ij}\right]} \quad =$$

$$= \frac{\exp\left[Q(g_{ij}=1,g_{-ij},x;\theta)-Q(g_{ij}=0,g_{-ij},x;\theta)\right]}{1+\exp\left[Q(g_{ij}=1,g_{-ij},x;\theta)-Q(g_{ij}=0,g_{-ij},x;\theta)\right]} \tag{9}$$

$$= \frac{\exp\left[u(x_i,x_j;\alpha)+u(x_j,x_i;\alpha)+\beta\sum_{r\neq i,j}^{n}(g_{jr}+g_{ir})+\gamma\sum_{r\neq i,j}^{n}(g_{jr}g_{ri}+g_{ir}g_{rj})\right]}{1+\exp\left[u(x_i,x_j;\alpha)+u(x_j,x_i;\alpha)+\beta\sum_{r\neq i,j}^{n}(g_{jr}+g_{ir})+\gamma\sum_{r\neq i,j}^{n}(g_{jr}g_{ri}+g_{ir}g_{rj})\right]} \tag{10}$$

Therefore, the potential fully characterizes the conditional choice probabilities of the firms.

The sequence of graphs generated by the network formation game is a Markov chain (Levin et al. (2008), Meyn and Tweedie (2009)), converging to a unique stationary equilibrium distribution over networks, which we can characterize in closed-form as

$$\pi(g,x;\theta)=\frac{\exp\left[Q(g,x;\theta)\right]}{c(\theta,x)} \tag{11}$$

where $Q(g,x;\theta)$ is the potential function in (5) and $c(\theta,x)$ is a normalizing constant that sums over all possible networks with n nodes (we call this set $\mathcal{G}$)

$$c(\theta,x)=\sum_{\omega\in\mathcal{G}}\exp\left[Q(\omega,x;\theta)\right] \tag{12}$$

The normalizing constant (12) guarantees that equation (11) is a proper distribution, in other words, that it sums to one when we sum over all possible networks.[10]

Equation (11) is the likelihood of observing a particular network structure in the long run. Notice that this likelihood has peaked at the maxima of the potential function and that, therefore, in the long run, we are most likely to observe the networks with high potential. By the discussion above, we know that all the pairwise stable equilibrium networks can be found as maxima of the potential function; as a consequence, the model predicts that, in the long run, we observe equilibrium networks with very high probability.

To perform the estimation, we will assume that the network of firms observed in our data is a realization of the long-run stationary equilibrium of the model. Accordingly, the distribution (11)

[10]The proof of convergence and the computation of (11) can be found in Mele (2017*a*) and Mele and Zhu (2017).

is the likelihood of observing a particular network.

Finally, our model has a very important property. Let's consider the potential function (5) and its elements. The sum

$$t_S(g,x) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{r\neq i,j}^{n} g_{ij}g_{jr} \tag{13}$$

corresponds to the number of 2-stars in network g; while the sum

$$t_T(g,x) = \frac{1}{6}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{r\neq i,j}^{n} g_{ij}g_{jr}g_{ri} \tag{14}$$

is the number of triangles in network g. Thus the likelihood of the network formation game (11) corresponds to the likelihood of an ERGM with 2-stars and triangles. We can easily obtain alternative ERGM specifications by changing the structure of the payoffs in Assumption 2.

3.4. **Extensions of the Network Formation Model.** While this exposition has focused on a dynamic model of network formation, characterizing the long-run distribution of networks in equilibrium, we can interpret this model in an alternative fashion. Consider a static network formation game, where each firm *simultaneously* chooses its portfolio of links. In such a game, multiple equilibrium networks are pairwise stable (with transfers). The stationary equilibrium of our model, described in the previous section, corresponds to a refinement of the static equilibrium, called *stochastic best-response dynamics* (Blume (1993), Jackson and Watts (2001)). According to this equilibrium refinement, firm pairs randomly encounter the opportunity to revise one of their link choices, but they make "mistakes" (modeled through random matching shocks ε_{ij}). The iteration of this stochastic best-response procedure generates the long-run distribution of networks in (11). This refinement of the equilibrium concept can also be considered an equilibrium selection device.

The network formation game presented in this paper is a model for a single network observation. We can estimate the parameters (α, β, γ) by observing only a single network. However, if we have multiple networks or we observe the network over time, we can exploit the information contained in the dynamics to identify richer sets of payoffs.

We can generalize the payoff functions used in this paper to include more externalities from link formation and to accommodate externalities that also depend on the observable characteristics of each firm. However, to precisely estimate such payoffs we need a larger network or multiple network observations.

4. Estimation, Data and Specification

We follow a Bayesian approach and estimate the posterior distribution of parameters, using the simulation methods developed in Mele (2017*a*) and Caimo and Friel (2011). The posterior is the conditional probability distribution of the parameters θ, given the observed network and our model. The prior distribution $p(\theta)$ summarizes prior knowledge regarding the parameters, and the posterior $p(\theta|g, x)$ is computed as

$$p(\theta|g,x) = \frac{\pi(g,x;\theta)p(\theta)}{\int_{\Theta} \pi(g,x;\eta)p(\eta)d\eta} \tag{15}$$

where $\pi(g, x; \theta)$ is the stationary equilibrium distribution (11) (that is, the likelihood of the observed network), and Θ is the parameter space.

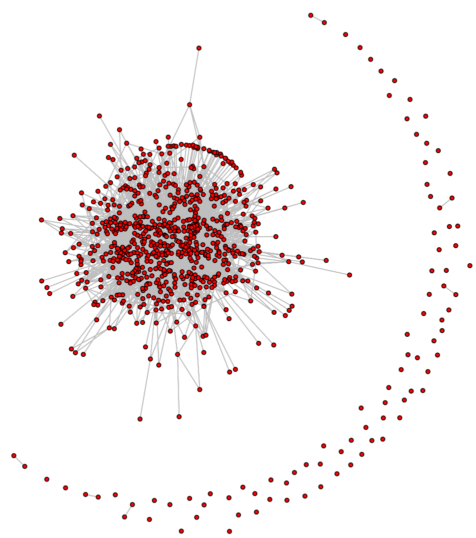
The Bayesian approach is well suited to estimating this class of models for several reasons. First, the literature on network estimation has established that some exponential random graphs models suffer from *degeneracy*. Degeneracy implies that the model puts the highest probability on a very small set of networks, usually the empty or full network. This is problematic because empirically it is very unusual to observe empty or complete networks. Researchers have dealt with this issue by providing alternative specifications of the model (Snijders (2002), Robins et al. (2007)), by including or excluding some terms (like triangles), and by "curving" the exponential distribution (Geyer (1992)). The Bayesian approach allows the researcher to specify priors that take into account how some parameter vectors imply degenerate models. A simple solution is to impose null prior probability for regions of the parameter space that generate almost empty or almost complete networks. A strength of our estimation strategy is that it does not rely on this extreme prior choice. Indeed, the estimation algorithm that we use in this paper, called the *exchange* algorithm (Murray et al. (2006), Mele (2017*a*), Caimo and Friel (2011)), samples from the posterior distribution of the parameters, giving more weight to those parameters that generate networks similar to the observed data. Therefore, the algorithm estimates a posterior where parameters that lead to degeneracy have small or null probability. This is a remarkable property that mitigates the issue of degeneracy explained above.

Second, we bypass the problem of computing the normalizing constant (12), because our algorithm recovers the posterior distribution of the parameters without evaluating the likelihood of the model. This is not a trivial point, because the constant is intractable (12) and cannot be computed in reasonable time even with a fast supercomputer, as explained in more detail below. The ERGM literature has developed alternative estimators, like the Markov Chain Monte Carlo Maximum Likelihood Estimator (MCMC-MLE) or the Maximum Pseudo-Likelihood Estimator (MPLE). The MCMC-MLE relies on a Monte Carlo approximation of the likelihood, through a long simulation of the model (Snijders (2002), Geyer and Thompson (1992)), which estimates the parameters by maximizing the approximated likelihood. As the number of simulations grows large, the parameters estimated with this method converge to the maximum likelihood estimate (Geyer and Thompson (1992)). However, practitioners have demonstrated several difficulties with these approximations, which may be poor if the chosen starting value of the simulation lies too far from the (exact) maximum likelihood estimate. In addition, there are cases in which the maximum likelihood estimate does not exist; when this is the case, the simulation output is unstable, leading to unreliable inference (Geyer and Thompson (1992), Snijders (2002), Butts (2009)). On the other hand, the Maximum Pseudo-Likelihood Estimator simplifies the estimation by considering the choice probabilities in (4) to be independent (Besag (1974), Wasserman and Pattison (1996)). As the network size grows, such estimates are consistent, though in many practical applications the estimates are poor and imprecise (Caimo and Friel (2011), Mele (2017*a*), Snijders (2002)). Furthermore, standard errors are usually underestimated and need adjustment.

4.1. **Application: Venture Capital Syndication.** We apply this methodology in the context of venture capital investments. Wilson (1968) defines a syndicate as a "group of individual decision makers who must make a common decision under uncertainty, and who, as a result, will receive a joint payoff to be shared among them". Venture capitalists operate in highly uncertain environments and tend to work alongside one another (Gu and Lu (2014), Lerner (1994)). These venture syndicates allow venture capitalists to deal with risks (Manigart et al. (2006)) in an uncertain environment. Also, syndication networks provide access to future investment opportunities (Hochberg et al. (2007)). Therefore, current syndication networks shape future co-investment ties (Zhang et al. (2017)).

Venture capital syndication is an apt empirical setting to examine network formation for two principal reasons. First, future investment ties rely on past ties (Zhang et al. (2017)). Second, organizations seek new partners based on firm attributes (Sorenson and Stuart (2008)). In sthis context, there is an interdependence among links, creating an estimation challenge that can be dealt with using our structural model and estimation method.

FIGURE 1. The network of venture capital co-investments in the medical device industry (1940-2013)



The network consists of 833 venture capital firms that have invested at least once in a company in the medical device industry between 1940 and 2013. There are 6997 co-investment links.

To this end, we built a dataset of all venture capital firms with headquarters in the United States that invested in the medical device industry between 1940 and 2013. Data on venture capitalists comes from the Venture Xpert database maintained by Thompson One. Our sample includes all the firms in this database that declare their investment type as "venture capital". This sample gives us a total of 833 firms. In our data, we have information, for each firm, about age, total capital under management, firm type, fund size and address of headquarters.

The network we examine in our empirical analysis is the co-investment network; a link connects two venture capital firms if they have invested together in a medical device firm. The resulting network is undirected, containing 833 nodes and 6997 links.

Our network data is shown in Figure 1. We observe a typical core-periphery structure, where some firms are extremely connected (the core), and some firms only form few links to the densely connected core (the periphery). We notice that some firms are isolated, as they did not form syndication with other firms during the period considered.

4.2. **Model specification.** We follow Mele (2017*a*) and specify a parsimoniuous model for the payoffs. The payoff from direct links has functional form

$$u(x_i, x_j; \alpha) = \alpha_0 + \sum_{p=1}^{P} \alpha_p f_p(x_i, x_j) \tag{16}$$

The first part of payoff (16) is α_0, which we interpret as the cost of forming a link. The second part of (16) is the sum of functions of i and j's attributes, that is $\sum_{p=1}^{P} \alpha_p f_p(x_i, x_j)$; this sum represents firm's i direct benefit of forming a link with firm j. Our implicit assumption is that the cost of forming links is the same across all the firms, but their benefits may vary according to their characteristics x_i and x_j.[11]

We have tried several specifications of the payoff function, and we will focus on the following in the empirical results

$$\begin{aligned} u(x_i, x_j; \alpha) &= \alpha_0 + \alpha_1 \mathbf{1}_{\{firmtype_i = firmtype_j\}} && (17)\\ &+ \alpha_2 |log(capital_i) - log(capital_j)| + \alpha_3 |age_i - age_j| \\ &+ \alpha_4 \mathbf{1}_{\{state_i = state_j\}} && (18) \end{aligned}$$

The variable $firmtype_i$ indicates whether the venture capital firm belongs to one of the following categories: {Angel Group, Bank Affiliated, Corporate PE/Venture, Endowment/Foundation or Pension Fund, Government Affiliated Program, Incubator/Development Program, Individuals, Insurance Firm Affiliate, Investment Management Firm, Private Equity Advisor or Fund of Funds,

[11]As explained in Mele (2017*b*) we can only identify the *net benefit* of a direct connection. An alternative specification could assume that the benefit of forming a link is constant and that the cost varies with observable characteristics.

Private Equity Firm, SBIC, Service Provider, University Program}. The variable $capital_i$ is the capital under management at venture capital firm i; age_i is the firm's age in 2013; and $state_i$ is the state where venture capital firm i has headquarters.

The model parameters to estimate are therefore $\theta = (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta, \gamma)$, and we denote the parameter space as Θ, that is $\theta \in \Theta$.

4.3. **Estimation Strategy.** To estimate the model we assume that the network we observe in the data is a stationary network, that is a draw from the stationary distribution of the theoretical model (11). In many applications the posterior distribution is available in closed-form; for example, it could be a normal distribution. However, in our case the posterior (15) is a *doubly intractable distribution* because it contains two intractable normalizing constants. Indeed, we can rewrite the posterior as

$$p(\theta|g,x) = \frac{\frac{\exp[Q(g,x;\theta)]}{c(\theta,x)}p(\theta)}{\int_\Theta \frac{\exp[Q(g,x;\eta)]}{c(\eta,x)}p(\eta)d\eta} = \frac{\frac{\exp[Q(g,x;\theta)]}{c(\theta,x)}p(\theta)}{p(g,x)} \tag{19}$$

The first intractable constant is $p(g,x) = \int_\Theta \frac{\exp[Q(g,x;\eta)]}{c(\eta,x)}p(\eta)d\eta$, which corresponds to the normalizing constant of the posterior distribution. This is also called *marginal likelihood* or *model evidence* in the Bayesian literature (Gelman et al. (2003)). The marginal likelihood is the probability of observing the network g given our specific model (that is, the ability of our model to explain the network in the data). This quantity is intractable because it involves an high-dimensional integration over the parameter space Θ. The computation of this constant is usually circumvented with the use of Metropolis-Hastings Markov Chain Monte Carlo (MCMC) algorithms (Liang et al. (2010)), simulating parameters from the posterior distribution in an iterative way. However, in this model we have an additional intractable normalizing constant in the likelihood

$$c(\theta,x) = \sum_{\omega\in\mathcal{G}} \exp\left[Q(\omega,x;\theta)\right]. \tag{20}$$

The sum is over all possible networks ω in the set $\mathcal{G}$, containing $2^{n(n-1)/2}$ networks. Even a small network with $n = 20$ firms would imply a sum over 2^{90} terms, each term involving the computation of the potential function $Q(\omega, x; \theta)$. A supercomputer will require days to compute such a constant at a specific parameter value. This implies that the likelihood and the posterior cannot be computed directly. With some algebra, we can show that first and second derivatives also depend on the normalizing constant $c(\theta, x)$, thus making maximum likelihood estimation quite challenging. This is an additional reason to prefer the Bayesian approach to the frequentist approach for this class of models.

We use an approximate exchange algorithm that circumvents the need to compute both constants $p(g, x)$ and $c(\theta, x)$ in the simulations. Murray et al. (2006) developed the original algorithm, later adapted to networks in Caimo and Friel (2011) and Mele (2017*a*).

For ease of exposition, we provide an informal description of the algorithm and its useful properties. The readers interested in technical details can find a formal description in our appendix and the proofs of convergence in Appendix B of Mele (2017*a*).

We start the simulation from parameter vector $\theta^{(0)}$ and network observed in the data g. The simulations proceed according to the following steps:

STEP 1 propose a new vector of parameters θ';

STEP 2 given the proposed parameters θ', simulate the network formation process using the model for R steps; collect the network from last step g';

STEP 3 check if the simulated network g' is similar to the network in the data g;

- if simulated and observed network are similar enough, accept the proposed parameter with high probability;
- otherwise, reject the parameter;

STEP 4 Repeat steps 1 to 3 for S times.

This algorithm samples parameter vectors that are most likely to generate the data. Once a parameter vector is proposed in STEP 1, we simulate R networks from the model in STEP 2. We need to choose R large enough to be sure that the last simulated network g' is approximately drawn from the stationary distribution (11); we then compare this *simulated* network g' with the observed network g, using a likelihood ratio. Mele (2017*a*) (Appendix B) shows that using the likelihood ratio eliminates the constant $c(\theta, x)$ and speeds up computations, thus avoiding the main computational bottleneck. If the simulated and observed networks are similar, that is, their likelihood ratio is close to 1, then the parameter vector we have proposed is very likely to generate the network in our data and should receive high probability in the posterior. That signifies that we should accept this proposed parameter with high probability. Viceversa, if the simulated and observed networks look very different and the likelihood ratio is very different from 1, the parameter we have proposed is unlikely to generate our data and thus should receive low probability in the posterior. Accordingly, we would reject this parameter vector with high probability.

This simulation will converge to the posterior distribution of the parameters, provided that both the number of network simulations R and parameter simulations S are large. In setting the number of simulations, we follow the applied probability literature that has established the speed of convergence for this type of algorithm. Bhamidi et al. (2011) and Mele (2017*a*) recently provided guidance on choosing the length of network simulations. We follow their suggestions here.

Our algorithm has several useful properties. First, the simulation strategy attenuates degeneracy problems, because parameters that generate almost-empty or almost-complete networks will receive very low or null probability in the estimated posterior. Such parameters will not be able to generate simulated networks that resemble our network data, thus the algorithm will reject them in favor of parameter vectors that generate networks similar to the observed data.

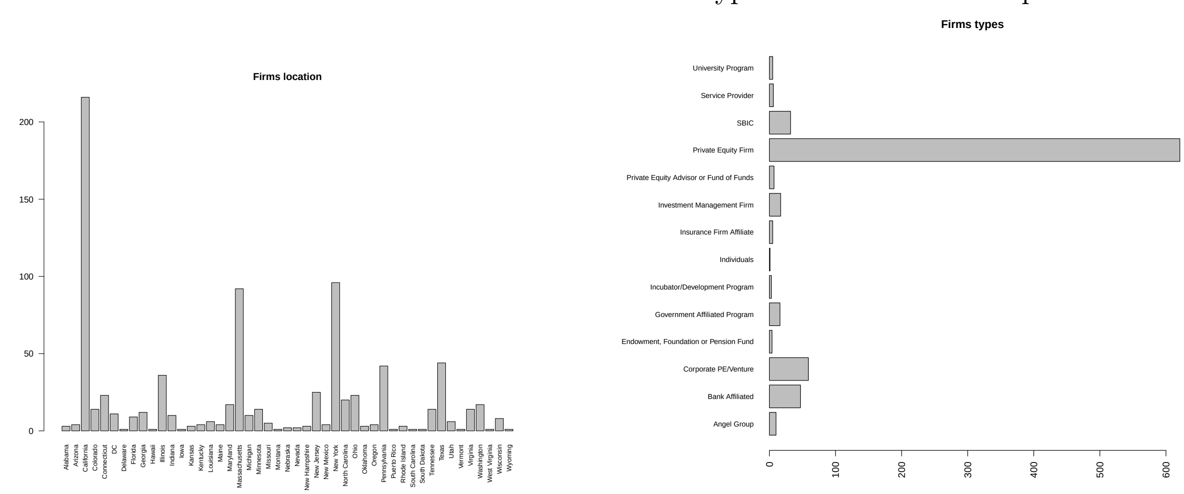
Second, our estimation procedure does not require a careful choice of starting value, unkike in the MCMC-MLE method. Indeed, the exchange algorithm converges to the correct posterior distribution independently of the starting value.[12] By contrast the MCMC-MLE is very sensitive to the initial parameter value, sometimes providing very poor approximations to the likelihood.

Third, the software for estimation is open source and available in the statistical package R. We use the package `Bergm`, developed by Caimo and Friel (2011) to perform our analysis.[13]

5. Empirical Results

5.1. **Descriptive statistics.** Our network contains 833 venture capital firms and 6997 co-investment partnerships. As shown in Figure 2, most of our firms are located in California, Massachussets and New York (left). Most companies in our data are private equity firms (right).

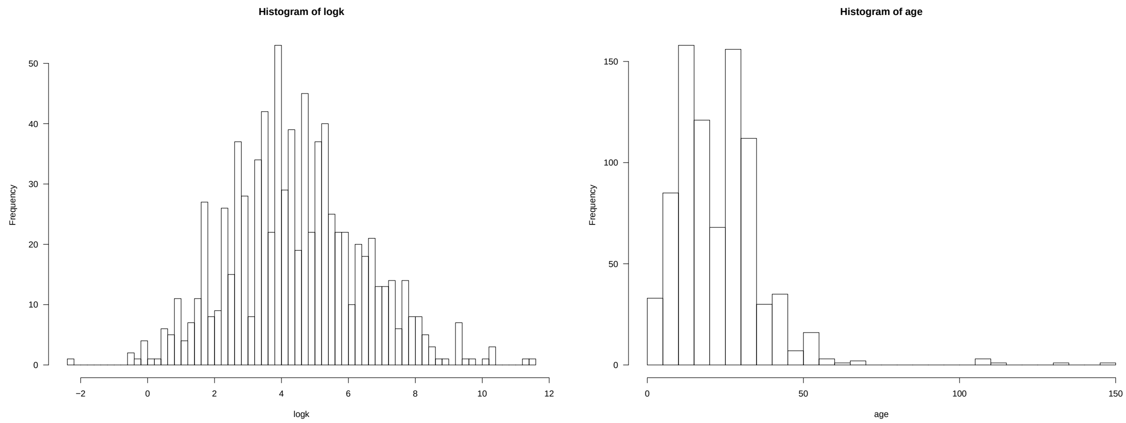
Figure 2. Distribution of location and types of firms in the sample



We show in Figure 3 the distribution of capital managed by the firm in logs (left) and the age of the company (right). The largest company manages a capital of 97,700 million dollars. Most companies are relatively young, however there are a few with ages above 50 years.

[12] For a formal proof of convergence, see Appendix B of Mele (2017*a*).
[13] An alternative package for some specific models is `netnew`, developed by Mele (2017*a*).

FIGURE 3. Distribution of capital managed (in logs) and age in the sample



The average firm has 34 links, while the median is 14. This substantial difference between mean and median is typical of core-periphery networks. In Figure 4 we show the whole degree distribution. The histogram shows the typical pattern of core-periphery networks; a few firms have most connections (core firms), while most companies have very few links to the core of the network (periphery firms). In the core, we find firms like Delphi Ventures, Johnson & Johnson Development Corp and Cdib Venture Management.

The next section provides estimates of structural and economically relevant parameters.

5.2. **Estimated structural parameters.** The posterior estimates for the model appear in Tables 1 and 2. We compare the specification with endogenous network effects (Table 1) to a standard logit model (Table 2) that does not include any endogenous network term. While the estimated posterior distributions are similar, we show below that the fit of the models is quite different.

In Table 1, the first column reports the posterior mean, while the second column is the standard deviation of the posterior. The third column shows the standard error of the posterior mean estimate. The fourth and fifth columns show the 95% credible interval. Credible intervals represent the possible values that the model's parameters can take given the data observed, for a given significance level (usually 95%).

FIGURE 4. Degree distribution of the co-investment network

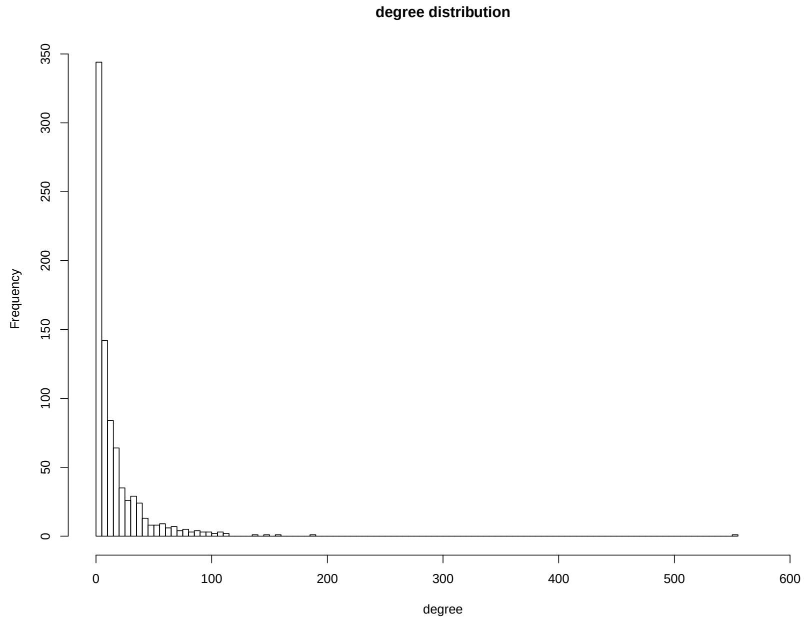


TABLE 1. Model with endogenous network variables

Variable	Posterior Mean	Posterior Std. Dev.	Standard Error	Credible Interval 2.5	97.5
Cost	-5.897	0.096	0.0001	-6.083	-5.709
Number of partners	0.006	0.0006	0.000006	0.005	0.007
Common partners	0.526	0.016	0.0001	0.494	0.558
Same firm type	-0.054	0.074	0.0008	-0.200	0.090
Abs. Difference Capital (log)	-0.023	0.021	0.0002	-0.064	0.018
Abs. Difference Age	-0.017	0.003	0.00003	-0.023	-0.012
Same State	0.641	0.096	0.003	0.452	0.829

Acceptance rate is 0.132

In Table 1 the estimated cost of forming a link is negative, as expected, and it appears quite high, with a posterior mean of -5.987. This result implies that syndication links are costly to maintain. The 95% credible interval does not include zero; therefore, we can credibly state that this estimate is negative. The estimate is precise, as shown by the small standard deviation of the posterior and the small standard error of the posterior mean. The economic interpretation is that forming an additional link will cost an average decrease of 5.897 in payoffs.

The estimated cost seems relatively high, so we need to consider the benefits of partnerships. These estimates are in the rest of Table 1, and include both endogenous network benefits (number of partners and shared partners) and payoffs from exogenous attributes. The first endogenous network term is the number of partners, which is determined by the number of 2-stars in the network. The credible interval for this term is positive, with a posterior mean of 0.006. The economic implication is that forming a link to a partner with an additional link would increase payoff by 0.006, on average.

The second endogenous network is the number of partners shared by the two companies. This term has quite a substantial effect on the payoff; an additional common partner would increase payoffs by an average of 0.526 (credible interval positive). This is a large effect, which drives the clustering observed in our network of venture capital firms and helps to explain the core-periphery structure.

The remaining variables in Table 1 show strong evidence of homophily for two firm attributes. Two venture capital firms belonging to the same firm type do not seem to have a higher-than-random probability of forming a link. This is shown by the credible interval that includes zero and estimates a quite large posterior probability around zero. In a standard frequentist analysis, this would correspond to saying that the coefficient is not significant.

One would expect that the amount of capital managed by the venture capital firm plays a vital role in determining link formation. Indeed our estimate shows that a 1% increase in the difference (in absolute value) of partners' capital decreases payoffs by 0.023 on average.[14] However, the credible interval includes zero, and therefore this conclusion is weakly supported.

There is evidence of homophily in the age of firms. The estimated posterior for the difference in age shows that a difference of 1 year decreases payoffs by 0.017, providing support to age homophily.

[14]The difference in capital managed by the venture capital firms is logged to facilitate interpretation.

Table 2. Model with no endogenous network variables

Variable	Posterior Mean	Posterior Std. Dev.	Standard Error	Credible Interval 2.5	97.5
Cost	-3.756	0.048	0.0004	-3.849	-3.661
Same firm type	0.009	0.041	0.0003	-0.070	0.089
Abs. Difference Capital (log)	-0.035	0.012	0.0001	-0.059	-0.012
Abs. Difference Age	-0.018	0.002	0.00002	-0.022	-0.014
Same State	1.049	0.059	0.0005	0.933	1.166

Acceptance rate is 0.189

The location is critical. Payoffs increase by 0.641 on average if partnered venture capital firms are in the same state.

In summary, venture capital firms tend to prefer syndication with firms of a different type, a similar level of managed capital, and similar age. There is a strong preference for firms in the same state. However, the homophily effects for type of firm and capital managed by the firm are not very precisely estimated, so these effects are ambiguous. Focusing on the structural network terms, we see a preference for syndication with firms with a higher number of partners and shared partners. This is most likely due to a reputation effect. Firms with many partners signal that they have had successful syndication in the past; firms with shared partners provide a screening device, as the common partner can certify the quality of the previous syndication.

In Table 2 we report the posterior estimates for the standard logit model, where the endogenous structural network terms are excluded from the specification. While most of the effects have a similar sign, we notice that the estimated cost is lower and the coefficient for homophily by the state is higher. This means that the equilibrium network effects for popularity and transitivity estimated in 1 are imputed to the homophily, rather than clustering. It is therefore important to determine which specification provides the best fit for our data because this will determine whether the core-periphery structure of the observed network is generated by pure homophily effects or by the transitivity and reputation effects implied by the effect of clustering in the payoff function.

5.3. **Goodness of Fit.** In the alliance literature, researchers have few ways to check the fit of a model. One of the advantages of our structural Bayesian approach is that we can easily check whether the estimated model posterior can replicate observed features of the network (see Caimo and Friel (2011), Mele (2017*a*), Caimo and Lomi (2015), Kim et al. (2015)).

To implement goodness-of-fit tests, we take a sample of 1000 parameter vectors from the posterior distribution estimated in the previous section. For each of these parameter vectors, we simulate our model and generate a network. We then compare the simulated networks to the observed network. A good fit should generate networks that are similar to the one observed in the data, for example, in terms of similar degree distribution and similar triangle counts.

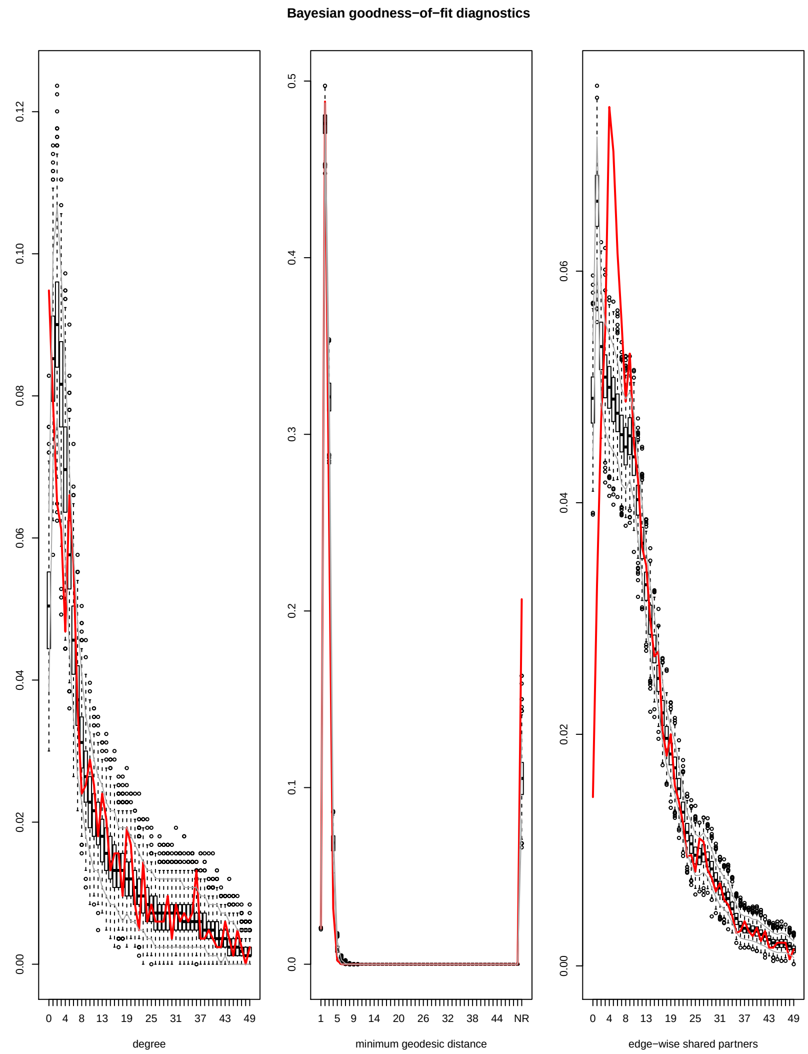
To make our tests more concrete, we compare the distribution of our simulated networks and the observed one with respect to features not included in the estimation. In our payoff specification, we explicitly include transitivity and popularity effects. Showing that our estimated model can replicate the transitivity and popularity observed in the network would not be surprising nor proof of good fit because we are targeting these network statistics directly in our specification.

Therefore, we provide tests for network statistics that are endogenous, but not included in the payoffs. The geodesic distance is the minimum shortest path in the network between two nodes. The number of edge-wise shared partners counts the number of pairs of nodes that have exactly k common neighbors, where k varies from 1 to the maximum number of nodes. Our goodness of fit test will focus on these statistics.[15]

We show the results in Figure 5 for the model estimated in Table 1. This figure shows three goodness-of-fit statistics: the degree distribution, the geodesic distance and the edge-wise shared partners. The red line represents the observed data. The boxplot is drawn at the interquartile range (25% and 75%). The grey lines represent 95% confidence bands. If our estimated model can

[15]This is also the default output of the goodness of fit test in the R package `Bergm` that we use for our empirical analysis.

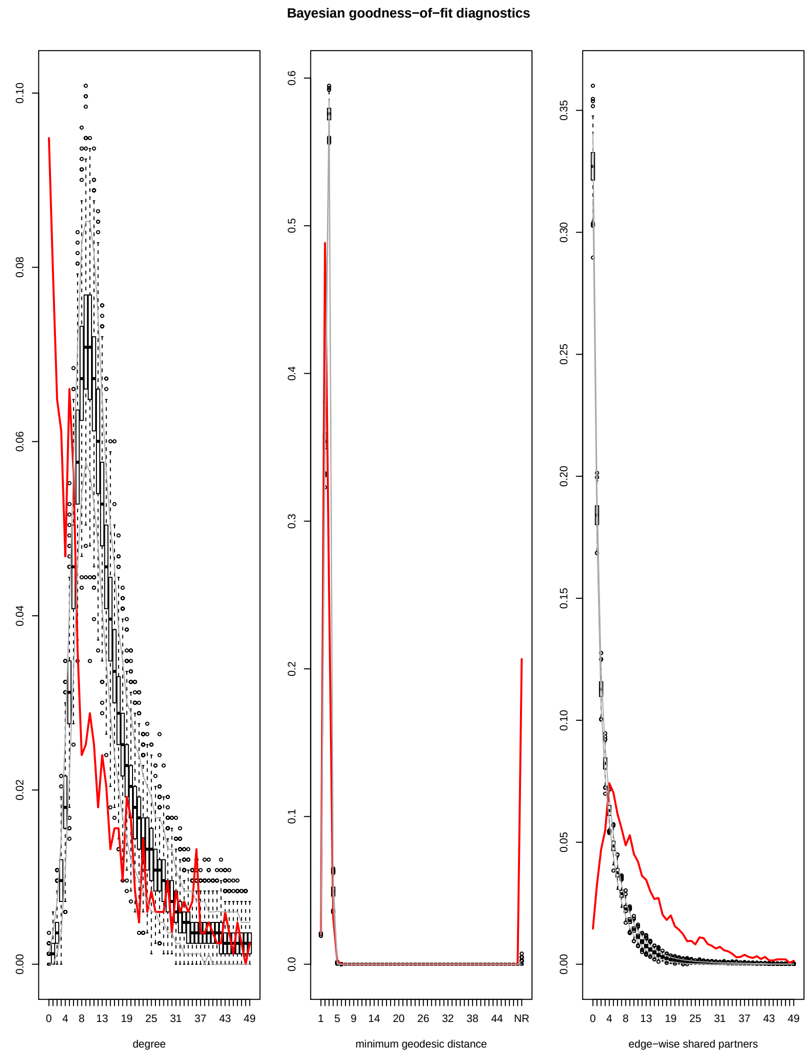
FIGURE 5. Model with externalities



replicate the observed network well, then the simulated networks should have degree distributions that are "close" to the observed degree distribution. In the picture, we expect to see most of the simulations falling within the boxplots, and certainly within the 95% bands. This is the case in Figure 1 for degree distribution, geodesic distance and edge-wise shared partners.

We compare these test results to the goodness-of-fit test for the model estimated in Table 2. Remember that this model does not include the endogenous network statistics (stars and triangles), and it is equivalent to a logistic regression model. The fit of this specification is inferior, as shown

FIGURE 6. model without network effects



in Figure 6. This implies that it is essential to include the endogenous network effects in the payoff functions.

5.4. **Counterfactual policy experiments.** The main advantage of a structural model with respect to reduced-form models is the possibility to run counterfactual experiments. For example, what would be the effect of several firms entering the market in a particular period on the equilibrium network?

Answering this question using a reduced-form approach will incur in the so-called Lucas Critique (Lucas (1976)). This is because a logit or probit model can be considered a partial equilibrium model, where the general equilibrium effects of a policy change are ignored. In practice, the parameters estimated using a reduced-form approach, are not policy-invariant, therefore cannot be used to make predictions on the effect of policy changes. Our equilibrium model, on the other hand, includes the equilibrium feedback directly in the utility function and models the equilibrium in the strategic network formation game.

An additional advantage of our structural model is that quantities like centrality, clustering and density of the network are considered equilibrium quantities. When there is a policy change, we expect those network statistics to change as well. This will impact how our firms make decisions about alliances and links, therefore further impacting the value of centrality, clustering and density. This feedback effect generated by the equilibrium model is absent in a standard logit or probit model.

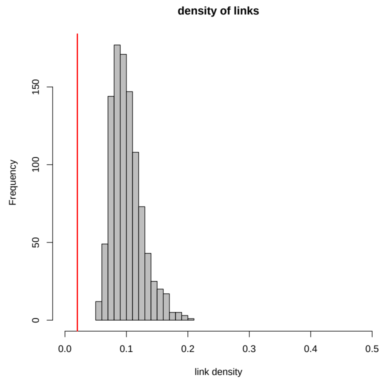
We use our estimated model to see how the equilibrium networks would change in three different scenarios: 1) Entry of 10 Venture capital firms in New York; 2) Entry of 5 firms in California; 3) Regulation that requires a minimum capital to enter the market.[16]

For each policy experiment, we simulate 1000 equilibrium networks using the posterior esimates of the structural parameters.[17] For each of these networks we compute density, clustering, homophily by firm type and state, and the degree distributions. We compare each of these structural equilibrium features with the observed network. This comparison allows us to determine the effect of a shock or policy change to the equilibrium network architecture.

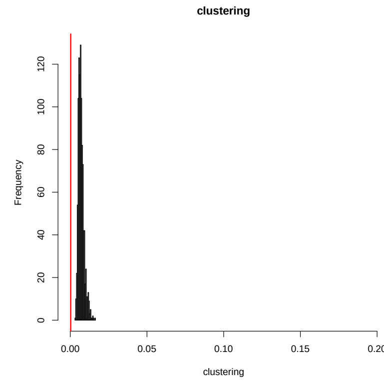
[16]In the last couterfactual policy simulation, we assume that firms that do not fulfill the minimum capital requirement exit the market. We assume that firms do not merge to avoid the consequences of the new policy, because the decision to merge is not modeled explicitly in our framework. In principle, we could simulate the merging decisions using an auxiliary model, but we keep things simple in our example.

[17]Each network is simulated using a large number of steps, to make sure that the network is an approximate draw from the new equilibrium of the model, after the policy change.

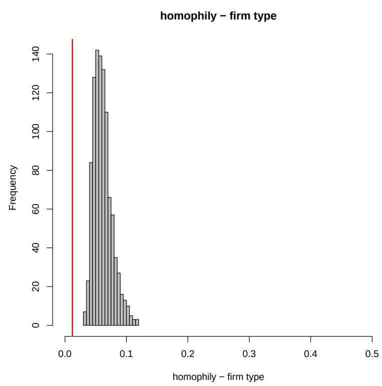
Figure 7. Policy counterfactual: Entry of 10 firms in New York



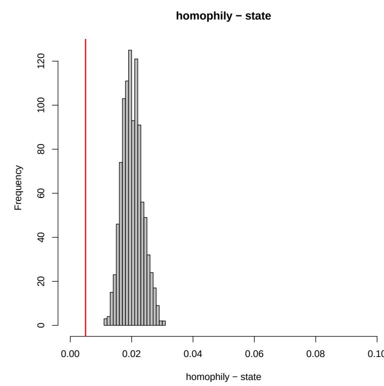
(A) Link density



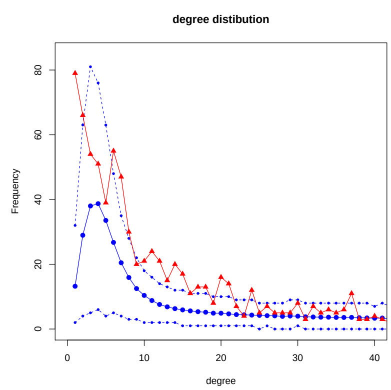
(B) Clustering



(C) Homophily - Firm type



(D) Homophily - State



(E) Degree distribution

The red line represents the observed network feature. The histograms are obtained from 1000 network simulations from the posterior distribution.

In Figure 7 we show the results of 1000 simulations for the entry of 10 venture capital firms in the New York area. In the simulation all the new firms entering the New York market are assumed to be Private Equity Firms; the (log) capital managed by the entrants is randomly drawn from a normal distribution with mean and standard deviation equal to the observed (log) capital mean and standard deviation.[18] Therefore, the entrants are on average similar to the incumbent firms.

We focus on five structural features of the new equilibrium network: (A) the density of links,[19] (B) the clustering,[20] (C) the level of homophily by firm type, (D) the level of homophily by state, and (E) the degree distribution. In panels (A) through (D) in Figure 7, the red lines represent the values for the network observed in our data, while the histograms show the distribution of the simulated equilibrium networks after the entry of the new firms. The blue lines in panel (E) represent the average degree distribution over 1000 simulations and the 2.5% and 97.5% quantiles, while the red line represent the observed degree distribution in the network data.

Our results show that when these 10 firms enter the New York Market, there is a general equilibrium effect that increases both density and clustering of the network (on average). Furthermore our simulations show that the density and clustering of the network are significantly different than the observed one. The equilibrium level of Homophily is also increased, both for state and firm type. The degree distribution also changes as a consequence of the new entries in the market. However, we notice that the firms that used to form fewer links before the arrival of the new entrants, tend to form on average even fewer links; this is counterbalanced by the high variance shown in our simulations for firms with low degree. On the other hand, firms that had a relatively higher degree, seem to maintain their average degree. Our simulations also show a decrease of the firms that are

[18] This allows the entrants to be quite similar to the existing firms, in terms of capital endowments.

[19] We measure the ratio of total number of links over the maximum number of links.

[20] Our measure of clustering is the total number of triangles divided by $n(n-1)(n-2)/6$. See Jackson (2008) for alternative ways to measure clustering.

completely isolated, thus contributing to the aggregate increase in density of the new equilibrium networks.

In Figure 8 we show a counterfactual where 5 venture capital firms enter the market in California. The results are quite similar to the entry of 10 firms in New York.

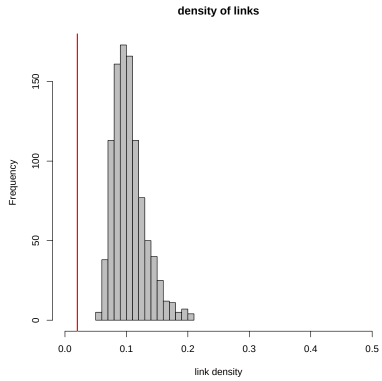
The final counterfactual is a policy in which the regulator imposes a minimum capital requirement for venture capital firms. In our simulation we impose a minimum capital that is the observed 25% quantile of capital managed by the firms in the market. To keep the simulation simple we assume that the firms that do not have enough capital will exit the market. We therefore abstract from the possibility of mergers to fulfill the new capital requirements. The main reason is that we have not modeled the merger decision in our model, and therefore this would not follow the spirit of the structural analysis.

The results of this counterfactual are shown in Figure 9. After the policy, there are only 625 firms in the market with the minimum required capital. The new equilibrium configuration involves a denser network and a much more clustered network. These effects are much larger than in the previous counterfactuals. In addition, the analysis of Panel (E) shows that the degree distribution is quite different after the implementation of the minimum capital requirement. Most of the change is concentrated in the set of firms with no links before the policy change. In the new equilibrium the number of firms with no syndication links drops on average by 70. This contributes to the increase in density. It is not surprising that the degree distribution shifts down, as there are fewer firms in the market; however, the shape of the degree distribution significantly changes.

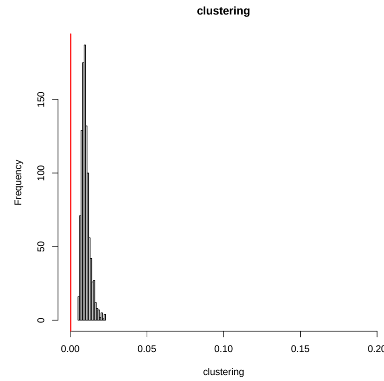
6. Conclusions and Discussion

Prior studies have focused on how firms forge ties based on their resource needs being fulfilled by partners (Chung et al. (2000), Rothaermel and Boeker (2008)), and trust due to repeated ties

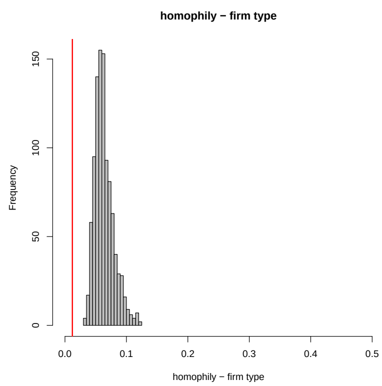
FIGURE 8. Policy counterfactual: Entry of 5 firms in California



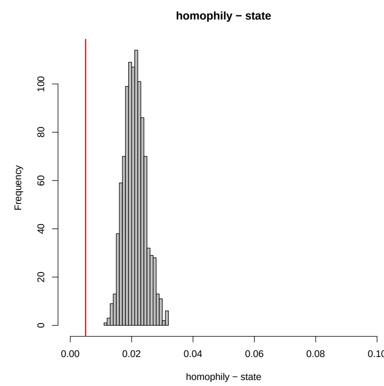
(A) Link density



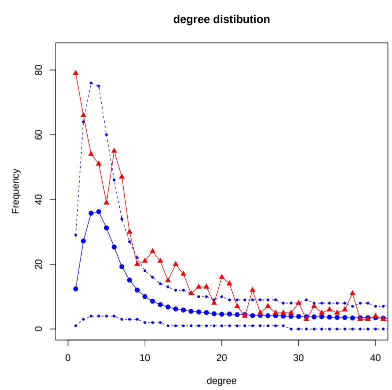
(B) Clustering



(C) Homophily - Firm type



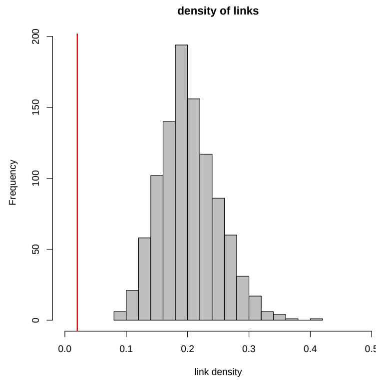
(D) Homophily - State



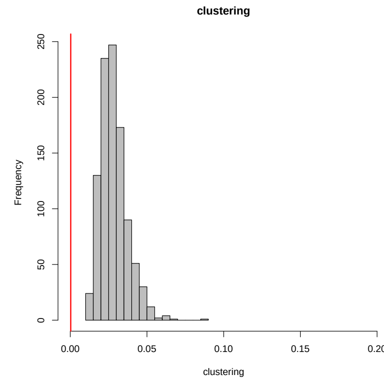
(E) Degree distribution

The red line represents the observed network feature. The histograms are obtained from 1000 network simulations from the posterior distribution.

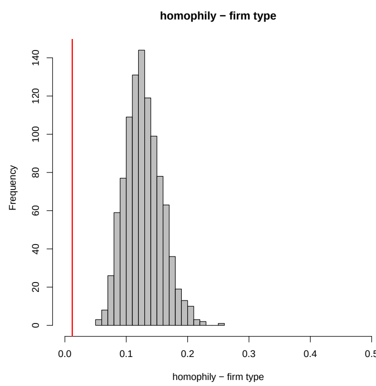
Figure 9. Policy counterfactual: Minimum capital requirement



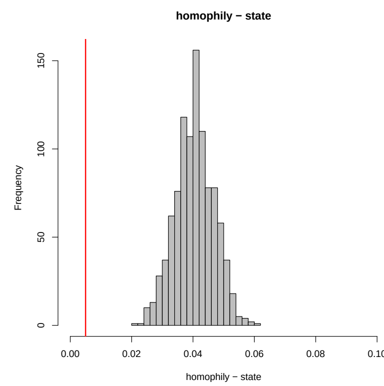
(A) Link density



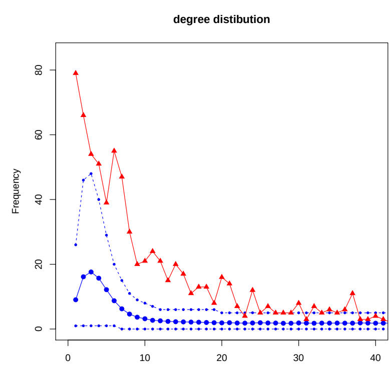
(B) Clustering



(C) Homophily - Firm type



(D) Homophily - State



(E) Degree distribution

The red line represents the observed network feature. The histograms are obtained from 1000 network simulations from the posterior distribution.

(Gulati (1995b),Gulati and Gargiulo (1999)). The most salient challenge to progress in this area is a lack of methods to control for the endogenous nature of network formation stemming from interdependence among ties. A reduced form estimation using the binary choice model to understand link formation is lacking in its ability to address the endogenous nature of tie formation. More recently, Exponential Random Graph Models or ERGMs (Snijders (2002)) have allowed researchers to tackle this challenge (Kim et al. (2015)). Exponential Random Graph Models have aided the study of network formation. However, ERGM'S suffer from two significant challenges. First, it is difficult to draw economic meaning from the estimated parameters. Second, this modeling strategy suffers from degeneracy problems, where the distribution tends to an empty or complete network.

In this paper, we develop a strategic model of network formation that transparently identify firms' incentives to form links in equilibrium. This approach controls for the endogeneity of network effects, such as transitivity, by imposing strategic equilibrium conditions on the linking decisions. As a consequence, we obtain a coherent economic and statistical framework to estimate the effect of network terms, taking into account their endogeneity. Furthermore, our Bayesian estimation strategy improves existing methods by decreasing degeneracy problems and guaranteeing convergence to the posterior distribution of the structural parameters. The estimated parameters correspond to meaningful economic concepts, such as the net benefit of syndication; and the network structural pieces such as density, clustering or homophily are equilibrium quantities. One of the key features of this approach is that it allows us to perform analysis for policy counter-factuals. Examining policy experiments and its impact on the network structure allows us to decode the "what if" questions. These counter-factual experiments allow managers or organizations to understand the implication of their strategic decisions on the emergence of inter-organizational network structure.

We demonstrate the significance of using structural models to study tie formation by applying our model to co-investment networks of venture capital firms. We focus on the venture capital firms that have invested in the medical device industry. A visualization of the data shows a core periphery structure. The structural model is used to examine the emergence of this network structure and a bayesian ERGM approach is used to estimate the parameters. The results shows that the venture capital firms form syndicates as result of their preference of firms that are similar to them regarding age, managed capital, and geographic location. These results are inline with the extant literature on tie formation among venture capital firms(Gulati and Gargiulo (1999), Sorenson and Stuart. (2001), Kogut et al. (2007)). However, the structural model adds to this conversation on venture capital syndication by allowing researchers to examine policy experiments. In this study, we examine two such policy counterfactuals — how the entry of firms or the imposition of a minimum level of managed capital could change the syndication network or network structure? Our simulations show that in response to these shocks, the network becomes more clustered and dense; firms that had no links before the shock, will form at least a tie, thus increasing density. These new ties are to similar firms, thus increasing the levels of homophily in the network.

This approach can be adopted in management research to study other kinds of clustering in networks. A similar core-periphery structure in a network graph was recorded in different contexts in management literature. For example, Fleming et al. (2004) examines the emergence of a dense cluster in the network of patents' co-authors in the early 1990s in Silicon Valley. This model would be applicable to study the emergence of co-author networks, and predict changes in the network structure based on competition or any other policy change. Another context that our model could be used is to provide a framework to understand how the endogenous ties create knowledge spillovers in equilibrium. Furthermore, our approach allows the simulation of counterfactuals; how public policies affect knowledge diffusion? What is the effect of subsidies to research and innovation on the structure of the network?

While our method has several advantages listed in the main text, we need to acknowledge some minor limitations. First, estimation is computationally intensive. In particular, the simulation of the posterior distribution may be more computationally intensive than the commonly used Markov Chain Monte Carlo Maximum Likelihood Estimator (MCMC-MLE). We note that this problem will decrease in a few years with faster hardware. Second, we model payoffs based on observable firms' characteristics, though it is possible that unobserved heterogeneity plays a role in the formation of links. Introducing unobserved heterogeneity in our model is possible at the cost of a substantial increase in computational burden. While this is an important direction of research, we leave the detail of this extension to future work.

References

Ahuja, G. (2000), 'The duality of collaboration: Inducements and opportunities in the formation of interfirm linkages', *Strategic Management Journal* **21**, 317–343.

Ahuja, Gautam, Francisco Polidoro and Will Mitchell (2009), 'Structural homophily or social asymmetry? the formation of alliances by poorly embedded firms', *Strategic Management JOURNAL* **30**(9), 941–958.

Badev, Anton (2013), Discrete games in endogenous networks: Theory and policy. UPenn.

Bala, Venkatesh and Sanjeev Goyal (2000), 'A noncooperative model of network formation', *Econometrica* **68**(5), 1181–1229.

Beckman, C. M., P. R. Haunschild and D. J. Phillips (2004), 'Friends or strangers? firm-specific uncertainty, market uncertainty, and network partner selection', *Organization Science* **15**, 259–275.

Besag, Julian (1974), 'Spatial interaction and the statistical analysis od lattice systems', *Journal of the Royal Statistical Society Series B (Methodological)* **36**(2), 192–236.

Bhamidi, Shankar, Guy Bresler and Allan Sly (2011), 'Mixing time of exponential random graphs', *The Annals of Applied Probability* **21**(6), 2146–2170.

Blume, Lawrence E. (1993), 'The statistical mechanics of strategic interaction', *Games and Economic Behavior* **5**(3), 387–424.

Burt, R. S. (1992), *Structural Holes*, Harvand University Press, Cambridge MA.

Butts, Carter (2009), Using potential games to parameterize erg models. working paper.

Caimo, Alberto and Alessandro Lomi (2015), 'Knowledge sharing in organizations: A bayesian analysis of the role of reciprocity and formal structure', *Journal of Management* **41**(2), 665–691.

Caimo, Alberto and Nial Friel (2011), 'Bayesian inference for exponential random graph models', *Social Networks* **33**(1), 41–55.

Chandrasekhar, Arun and Matthew Jackson (2014), Tractable and consistent exponential random graph models. Stanford University.

Chung, Seungwha, Harbir Singh and Kyungmook Lee (2000), 'Complementarity, status similarity and social capital as drivers of alliance formation', *Strategic management JOURNAL* pp. 1–22.

Conley, Timothy G. (1999), 'Gmm estimation with cross sectional dependence', *Journal of Econometrics* **92**(1), 1–45.

Conley, Timothy G. and Giorgio Topa (2007), 'Estimating dynamic local interactions models', *Journal of Econometrics* **140**(1), 282–303.

Currarini, Sergio, Matthew O. Jackson and Paolo Pin (2009), 'An economic model of friendship: Homophily, minorities, and segregation', *Econometrica* **77**(4), 1003–1045.

Currarini, Sergio, Matthew O. Jackson and Paolo Pin (2010), 'Identifying the roles of race-based choice and chance in high school friendship network formation', *the Proceedings of the National Academy of Sciences* **107**(11), 4857–4861.

Diaconis, Persi and Sourav Chatterjee (2013), 'Estimating and understanding exponential random graph models', *Annals of Statistics* **41**(5), 2428–2461.

Fleming, L., L. Colfer, A. Marin and J. McPhie. (2004), 'Why the valley went first: Agglomeration and emergence in regional inventor networks.'. working paper.

Fox, Jeremy, David Hsu and Chenyu Yang (forthcoming), 'Unobserved heterogeneity in matching games', *Journal of Political Economy* .

Garcia-Pont, Carlos and Nitin Nohria (2002), 'Local versus global mimetism: The dynamics of alliance formation in the automobile industry', *Strategic Management JOURNAL* **23**(4), 307–321.

Gelman, A., J. Carlin, H. Stern and D. Rubin (2003), *Bayesian Data Analysis, Second Edition*, Chapman & Hall/CRC.

Geyer, Charles and Elizabeth Thompson (1992), 'Constrained monte carlo maximum likelihood for depedendent data', *Journal of the Royal Statistical Society, Series B (Methodological)* **54**(3), 657–699.

Geyer, Charles J. (1992), 'Practical markov chain monte carlo', *Statistical Science* **7**, 473–511.

Ghosh, Anindya, Ram Ranganathan and Lori Rosenkopf (2016), 'The impact of context and model choice on the determinants of strategic alliance formation: Evidence from a staged replication study', *Strategic Management JOURNAL* **37**(11), 2204–2221.

Gu, Qian and Xiaohui Lu (2014), 'Unraveling the mechanisms of reputation and alliance formation: A study of venture capital syndication in china', *Strategic Management Journal* **35**(5), 739–750.

Gulati, R. (1995*a*), 'Does familiarity breed trust? the implications of repeated ties for contractual choice in alliances', *Academy of Management Journal* **38**, 85–112.

Gulati, R. (1995*b*), 'Social structure and alliance formation patterns: A longitudinal analysis', *Administrative Science Quarterly* **40**(619-652).

Gulati, R. (1999), 'Network location and learning: The influence of network resources and firm capabilities on alliance formation', *Strategic Management Journal* **20**(5), 397–420.

Gulati, R. and M. Gargiulo (1999), 'Where do interorganizational networks come from?', *American Journal of Sociology* **104**(5), 1439–1493.

Hallen, Benjamin L (2008), 'The causes and consequences of the initial network positions of new organizations: From whom do entrepreneurs receive investments?', *Administrative Science Quarterly* **53**(4), 685–718.

Heckman, James J. (1978), 'Dummy endogenous variables in a simultaneous equation system', *Econometrica* **46**(4), 931–959.

Hochberg, Yael V, Alexander Ljungqvist and Yang Lu (2007), 'Whom you know matters: Venture capital networks and investment performance', *The Journal of Finance* **62**(1), 251–301.

Jackson, Matthew and Alison Watts (2001), 'The existence of pairwise stable networks', *Seoul Journal of Economics* **14**(3), 299–321.

Jackson, Matthew and Asher Wolinsky (1996), 'A strategic model of social and economic networks', *Journal of Economic Theory* **71**(1), 44–74.

Jackson, Matthew O. (2008), *Social and Economics Networks*, Princeton.

Kim, J. Y., M. Howard, E. C. Pahnke and W. Boeker (2015), 'Understanding network formation in strategy research: Exponential random graph models', *Strategic Management Journal* **37**(1), 22–44.

Kogut, Bruce, Pietro Urso and Gordon Walker (2007), 'Emergent properties of a new financial market: American venture capital syndication,1960-2005', *Management Science* **53**(7), 1181–1198.

Konig, Michael, Xiaodong Liu and Yves Zenou (2017), R&d networks: theory, empirical and policy implications. unpublished manuscript.

König, Michael David (2016), 'The formation of networks with local spillovers and limited observability', *Theoretical Economics* **11**(3), 813–863.

Lerner, Joshua (1994), 'The syndication of venture capital investments', *Financial management* pp. 16–27.

Levin, David. A, Yuval Peres and Elizabeth L. Wilmer (2008), *Markov Chains and Mixing Times*, American Mathematical Society.

Li, Stan Xiao and Timothy J Rowley (2002), 'Inertia and evaluation mechanisms in interorganizational partner selection: Syndicate formation among us investment banks', *Academy of Management JOURNAL* **45**(6), 1104–1119.

Liang, Faming, Chuanhai Liu and Raymond J. Carroll (2010), *Advanced Markov Chain Monte Carlo Methods: Learning from Past Samples*, John Wiley and Sons, Ltd.

Lucas, Robert Jr. (1976), Econometric policy evaluation: A critique, *in* K.Brunner and A.Meltzer, eds, 'The Phillips Curve and Labor Markets', Vol. 1 of *Carnegie-Rochester Conference Series on Public Policy*, pp. 19–46.

Madhavan, R., D. R. Gnyawali and J. He (2004), 'Two's company, three's a crowd? triads in cooperative-competitive networks', *Academy of Management Journal* **47**, 918–927.

Manigart, Sophie, Andy Lockett, Miguel Meuleman, Mike Wright, Hans Landström, Hans Bruining, Philippe Desbrières and Ulrich Hommel (2006), 'Venture capitalists' decision to syndicate', *Entrepreneurship Theory and Practice* **30**(2), 131–153.

Marjoram, Paul, John Molitor, Vincent Plagnol and Simon Tavaré (2003), 'Markov chain monte carlo without likelihoods', *Proceedings of the National Academy of Sciences* **100**(26), 15324–15328.

Martin, Geoffrey, Remzi Gözübüyük and Manuel Becerra (2015), 'Interlocks and firm performance: The role of uncertainty in the directorate interlock-performance relationship', *Strategic Management JOURNAL* **36**(2), 235–253.

Mele, Angelo (2017*a*), 'A structural model of dense network formation', *Econometrica* **85**, 825–850.

Mele, Angelo (2017*b*), A structural model of homophily and clustering in social networks. working paper.

Mele, Angelo and Lingjiong Zhu (2017), Approximate variational estimation for a model of network formation. Johns Hopkins University.

Meyn, Sean and Richard L. Tweedie (2009), *Markov Chains and Stochastic Stability*, Cambridge University Press.

Mindruta, Denisa, Mahka Moeen and Rajshree Agarwal (2016), 'A two-sided matching approach for partner selection and assessing complementarities in partners' attributes in inter-firm alliances', *Strategic Management JOURNAL* **37**(1), 206–231.

Monderer, Dov and Lloyd Shapley (1996), 'Potential games', *Games and Economic Behavior* **14**, 124–143.

Murray, Iain A., Zoubin Ghahramani and David J. C. MacKay (2006), 'Mcmc for doubly-intractable distributions', *Uncertainty in Artificial Intelligence* .

Nakajima, Ryo (2007), 'Measuring peer effects on youth smoking behavior', *Review of Economic Studies* **74**(3), 897–935.

Podolny, J.M. (1993), 'A status-based model of market competition', *American Journal of Sociology* **98**(4), 829–872.

Powell, W. W., K. W. Koput and L. Smith-Doerr (1996), 'Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology', *Administrative Science Quarterly* **41**, 116–145.

Reuer, Jeffrey J and Nandini Lahiri (2013), 'Searching for alliance partners: Effects of geographic distance on the formation of r&d collaborations', *Organization Science* **25**(1), 283–298.

Robins, G., T. Snijders, P. Wang, M. Handcock and P. Pattison (2007), 'Recent developments in exponential random graph (p*) models for social networks', *Social Networks* .

Rothaermel, Frank T and Warren Boeker (2008), 'Old technology meets new technology: Complementarities, similarities, and alliance formation', *Strategic Management JOURNAL* **29**(1), 47–77.

Shipilov, Andrew V and Stan Xiao Li (2012), 'The missing link: The effect of customers on the formation of relationships among producers in the multiplex triads', *Organization Science* **23**(2), 472–491.

Snijders, Tom A.B (2002), 'Markov chain monte carlo estimation of exponential random graph models', *Journal of Social Structure* **3**(2).

Sorenson, O. and T. Stuart. (2001), 'Syndication networks and the spatial distribution of venture capital investments. amer', *American Journal of Sociology* **106**, 1546–1588.

Sorenson, O. and T. Stuart (2008), 'Bringing the context back in: settings and the search for syndicate partners in venture capital investment networks', *Administrative Science Quarterly* **53**, 266–298.

Stern, Ithai, Janet M Dukerich and Edward Zajac (2014), 'Unmixed signals: How reputation and status affect alliance formation', *Strategic Management JOURNAL* **35**(4), 512–531.

Wasserman, Stanley and Katherine Faust (1994), *Social Network Analysis: Methods and Applications*, Cambridge University Press.

Wasserman, Stanley and Philippa Pattison (1996), 'Logit models and logistic regressions for social networks: I. an introduction to markov graphs and p*', *Psychometrika* **61**(3), 401–425.

Zhang, L., Gupta A.K. and B.L. Hallen (2017), 'The conditional importance of prior ties: A group-level analysis of venture capital syndication', *Academy of Management Journal* **60**, 1360–1386.

Zhang, Lei, Anil K Gupta and Benjamin L Hallen (2017), 'The conditional importance of prior ties: A group-level analysis of venture capital syndication', *Academy of Management Journal* **60**(4), 1360–1386.

Appendix A. Estimation details

We estimated all the models using the package `Bergm` in R, developed by Caimo and Friel (2011). All the computations have been performed on a desktop Dell Precision T7620 with 2 Intel Xeon CPUs E5-2697 v2 with 12 Dual core processors at 2.7GHZ each and 64GB of RAM.

Each table of parameters estimates is obtained with 50000 parameters simulations after 5000 simulations are discarded as burn-in. For each parameter proposal, we simulate the model for 10000 iterations and pick the last simulated network to compute the acceptance ratio of the exchange algorithm. We use the snooker algorithm with 10 parallel simulations to improve convergence as implemented in the package (see Caimo and Friel (2011).

Codes for estimation are available upon request.

A.1. **Implementation of the exchange algorithm.** The exchange algorithm shown in the estimation section is computationally intensive and requires some tuning.

The algorithm proceeds as follows. At each iteration $s = 1, 2, ...$ and current parameter θ_s

(1) Propose a new parameter vector θ'

$$\theta' \sim q_\theta(\cdot|\theta_s)$$

(2) Given the proposed parameter θ', simulate a network g^* as follows. At iteration r

(a) Propose a new network g'

$$g' \sim q_g(\cdot|g_r)$$

(b) Then update the network at iteration $r+1$

$$g_{r+1} = \begin{cases} g' & \text{with prob. } \alpha_g \\ g_r & \text{with prob. } 1-\alpha_g \end{cases}$$

where α_g is

$$\alpha_g = \min\left\{1, \frac{\exp\left[Q(g', x; \theta')\right]}{\exp\left[Q(g_r, x; \theta')\right]} \frac{q_g(g_r|g')}{q_g(g'|g_r)}\right\}$$

(c) Iterate this process for $r = 1, ..., R$ and collect the last network $g_R = g^*$.

(3) Update the parameter at iteration $s+1$

$$\theta_{r+1} = \begin{cases} \theta' & \text{with prob. } \alpha_{ex} \\ \theta_r & \text{with prob. } 1-\alpha_{ex} \end{cases}$$

where the probability α_{ex} is given by

$$\alpha_{ex} = \min\left\{1, \frac{\exp\left[Q(g^*, x; \theta')\right]}{\exp\left[Q(g_r, x; \theta')\right]} \frac{\exp\left[Q(g_r, x; \theta_s)\right]}{\exp\left[Q(g^*, x; \theta_s)\right]} \frac{p(\theta')}{p(\theta_s)} \frac{q_\theta(\theta_s|\theta')}{q_\theta(\theta'|\theta_s)}\right\} \tag{21}$$

Notice that the probability (21) does not contain any normalizing constant. Neither κ nor $c(\theta, x)$ appear in the formulas; therefore our simulations are feasible. The simulations of parameters converge to the posterior distribution of the structural parameters. All the technical details and the proofs of convergence are in Appendix B of Mele (2017*a*).

The proposal distribution $q_\theta(\cdot|\theta_s)$ is a normal centered at the current parameter. To obtain the optimal variance for this proposal distribution, we estimate the model several times and adjust the variance of the proposal distribution to obtain better acceptance rates. The exchange algorithm has low acceptance rates compared to a standard Metropolis-Hastings algorithm. Our estimates have acceptance rates of 13% and 18%, compared with an optimal (asymptotic) rate of 25% for the Metropolis-Hastings.

The proposal distribution for networks $q_g(\cdot|g_r)$ selects a random pair of nodes and proposes to cut their link (if it exists) or form the link (if it does not exist). With a small probability, the proposed network swaps all the entries of the adjacency matrix. This step allows the sampler to reach other modes, being extremelty useful if the likelihood has multiple modes. For example, this is the case in several models analyzed in Mele (2017*a*).

Appendix B. Additional estimation results

This appendix reports alternative estimates and goodness of fit tests for the model. These results were obtained using Maximum Pseudolikelihood estimation and Monte Carlo Maximum Likelihood estimation. The MPLE results are shown in Table 3. Notice that this is a frequentist approach and it delivers estimated parameters and standard error. We notice that the estimates are quite different from the posterior means shown in Table 1, especially the parameters for Number of common partners.

Table 3. Model with endogenous network variables, estimated with Maximum Pseudolikelihood (MPLE)

Variable	Estimate	Std. Error
Cost	-5.078	0.039
Number of partners	0.007	0.0001
Common partners	0.268	0.002
Same firm type	-0.154	0.031
Abs. Difference Capital (log)	-0.016	0.009
Abs. Difference Age	-0.016	0.001
Same state	0.510	0.037

The main advantage of the MPLE is that it is quite fast and usually converges relatively fast to the final estimate. This is due to the fact that the Maximum Pseudolikelihood estimator assumes that the conditional choice probabilities of forming links factorize into the pseudolikelihood: the underlying assumption is that conditional on the rest of the network, the links are independent. This assumption is not satisfied if we believe that the actors (firms) are strategic in their decision to form links. There are theoretical results showing that MPLE estimates are consistent. On the other hand many practitioners complain of the poor performance of this estimator and the standard errors are imprecise and underestimated.

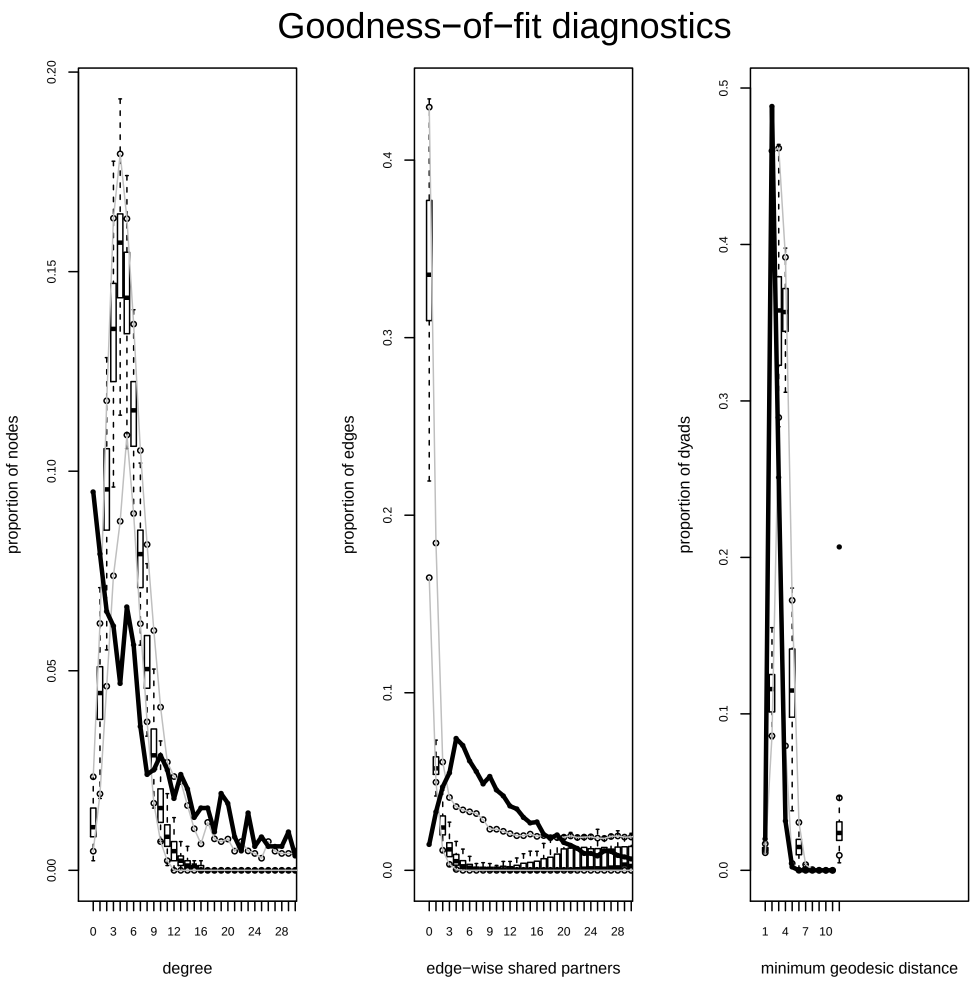
To give an idea of the poor performance of this estimator we provide goodness of fit tests similar to the ones performed after the Bayesian estimation. We report our tests in Figure 10. We simulate 100 networks from the model, using the parameters in Table 3 and compare the observed network to the distribution of the simulated ones. We focus on three network statistics: the degree distribution, the edge-wsie shared partners, and the geodesic distance. The solid line represents the observed values in our data, while the dotted lines are the 95% confidence levels for the simulated networks. If the solid line is contained within the dotted lines, then the estimated model is able to replicate the feature of the observed network.

The picture shown in Figure 10 provide evidence that the MPLE is inadequate to fit these data, and performs worse than our Bayesian approach.

Johns Hopkins University, Carey Business School, 100 International Dr, Baltimore, MD 21202

Johns Hopkins University, Carey Business School, 100 International Dr, Baltimore, MD 21202

FIGURE 10. Goodness of fit, model with network effects, Maximum Pseudolikelihood estimates



The black solid line represents the observed values of the degree distribution, distribution of edge-wise shared partners and minimu geodesic distance, respectively. The dotted lines represent the 95% confidence interval for the simulated model. The test is performed by simulating 100 networks from the equilibrium model. The parameters for the simulations are in Table 3. Each network is simulated with a MCMC run of 100000 steps.