A Structural Econometric Analysis of Network Formation Games Through Subnetworks*

Shuyang Sheng†

June 12, 2018

Abstract

The objective of this paper is to identify and estimate network formation models using observed data on network structure. We characterize network formation as a simultaneous-move game, where the utility from forming a link depends on the structure of the network, thereby generating strategic interactions between links. With the prevalence of multiple equilibria, the parameters are not necessarily point identified. We leave the equilibrium selection unrestricted and propose a partial identification approach. We derive bounds on the probability of observing a subnetwork, where a subnetwork is the restriction of a network to a subset of the individuals. Unlike the standard bounds as in Ciliberto and Tamer (2009), these subnetwork bounds are computationally tractable in large networks provided we consider small subnetworks. We provide Monte Carlo evidence that bounds from small subnetworks are informative in large networks.

JEL Classifications: C13, C31, C57, D85

KEYWORDS: Network formation, simultaneous-move games, multiple equilibria, subnetworks, partial identification, moment inequalities, simulation.

*This paper is a revision of Chapter 2 of my dissertation. I am very grateful to my advisor Geert Ridder for his enormously valuable advice and guidance. I also thank Aureo de Paula, Bryan Graham, Jinyong Hahn, Matthew Jackson, Rosa Matzkin, Roger Moon, Hashem Pesaran, Matthew Shum, Martin Weidner, Simon Wilkie, three anonymous referees, and seminar participants at USC, UCLA, UCSD, Johns Hopkins, Arizona, Pittsburgh, Tilburg, CORE, California Econometrics Conference, NASM, NAWM, EMES for helpful discussions and comments. Financial support from the USC Graduate School Dissertation Completion Fellowship is acknowledged. All the errors are mine.

†Department of Economics, UCLA, Los Angeles, CA 90095. E-mail: ssheng@econ.ucla.edu.

1 Introduction

Social and economic networks influence a variety of individual behaviors and outcomes, including educational achievement (Calvó-Armengol, Patacchini, and Zenou (2009)), employment (Calvó-Armengol and Jackson (2004)), technology adoption (Conley and Udry (2010)), consumption (Moretti (2011)), and smoking (Nakajima (2007)). As networks are often the result of individual decisions, understanding the formation of networks is important for the investigation of network effects. While the theoretical literature on network formation has flourished in the past decades (see Jackson (2008) and Goyal (2007) for a survey), econometric studies on the identification and estimation of network formation models are limited. The objective of this paper is to provide insight into this latter area. More precisely, assume that we observe the network structure, i.e., who is linked with whom. We propose new methods to identify and estimate the structural parameters in the model of network formation.

The statistical analysis of network formation dates back to the seminal work of Erdős and Rényi (1959), who proposed a random graph model where links are formed independently with a fixed probability. Statisticians later extended the Erdős-Rényi model to allow for dependence between links and developed a large class of exponential random graph models (ERGM) (e.g., Snijder (2002)). While ERGMs may well fit the observed network statistics, they usually lack microfoundations which are essential for counterfactual analysis. Alternatively, economists view network formation as the optimal choices of individuals who maximize their utilities. A simple and widely used empirical approach in this spirit is to employ a dyadic regression, where the formation of a link is modeled as a binary choice of the pair involved (e.g., Fafchamps and Gubert (2007), Mayer and Puller (2008)). In order to treat links in a network as independent observations, this approach needs to assume that there is no spillover from indirect friends (e.g., friends of friends), which could be restrictive in many applications given the prevalence of clustering (Jackson and Rogers (2007), Jackson, Barraquer and Tan (2012)). Graham (2017) extends dyadic regressions by allowing for individual fixed effects which can create interdependence between links. A more general class of network formation models permits utility externalities from indirect friends, thereby giving rise to strategic interactions between links (Christakis, Fowler, Imbens, and Kalyanaraman (2010), Mele (2017), Mele and Zhu (2017), Boucher and Mourifié

(2017), Miyauchi (2016), de Paula, Richards-Shubik and Tamer (2018), Leung (2015), Ridder and Sheng (2018), Leung (2018), Menzel (2017)). A contribution of this paper is to provide a different approach to the identification and estimation of such strategic network formation models.

A crucial problem in the identification of network formation models with strategic interactions is the presence of multiple equilibria. Boucher and Mourifié (2017) get around this problem by assuming there is a unique equilibrium in the observed data. Christakis et al. (2010) and Mele (2017) circumvent the multiplicity issue by considering a sequential model which typically yields a unique network or a unique stationary distribution over networks (Young (1993), Jackson and Watts (2002)). Unlike these studies, we admit multiple equilibria and do not impose restrictive assumptions on equilibrium selection. Since network outcomes may all belong to certain multiple equilibria, the parameters are not necessarily point identified. We propose a partial identification approach and examine what we can learn about the parameters from bounds on conditional choice probabilities.[1] The studies closest to ours are by Miyauchi (2016) and de Paula et al. (2018), who also consider bounds on conditional choice probabilities. They derive bounds under the restriction that either the utility externalities are nonnegative (Miyauchi, 2016) or agents can have a finite number of types and form a small number of links (de Paula et al., 2018). We consider bounds that hold for more general utility specifications.

The estimation of network formation models is typically computationally challenging because the number of possible networks is enormous: for n individuals the number of possible undirected networks is $2^{n(n-1)/2}$. In our case, the computation of the bounds may be intractable because it requires checking equilibrium conditions for all possible network configurations. We propose a completely new approach to tackle this computational problem. The idea is to make use of subnetworks. A subnetwork is the restriction of a network to a subset of the individuals. Under the equilibrium concept we consider (i.e., pairwise stability proposed by Jackson and Wolinsky (1996)), we can derive bounds on the probability of observing a subnetwork. These subnetwork bounds take into account not only the equilibrium of a subnetwork, but also the equilibrium of the rest of the network. Under our utility specification these

[1] Partial identification has been widely applied to game-theoretic models to cope with the identification issue due to multiple equilibria. See, e.g., Tamer (2003), Andrews, Berry and Jia (2004), Pakes, Porter, Ho and Ishii (2006), Berry and Tamer (2006), Ciliberto and Tamer (2009), Beresteanu, Molchanov, and Molinari (2011) among others.

subnetwork bounds are computationally tractable even in large networks as long as we only consider small subnetworks. This approach only needs choice probabilities within subnetworks, so it is still applicable if we do not observe an entire network, but links in subnetworks.

The estimation and inference of the identified set defined by the subnetwork inequalities is a straightforward application of the literature on partially identified models.[2] It is worth pointing out that our approach differs substantially from a recent strand of literature on large networks, which typically assumes that a single large network is observed (Leung (2015), Ridder and Sheng (2018), Leung (2018), Menzel (2017)). By assuming many networks, our approach does not need the restrictions that these studies have to impose to control for the dependence between links[3] and can be seen as complementary to them. In addition, we propose a novel method to compute the subnetwork bounds by transforming the search over equilibrium networks with a certain subnetwork into constraint optimization problems. This computational method is feasible in large networks as long as subnetworks are small.

The remainder of the paper is organized as follows. Section 2 develops the model. Section 3 addresses the problem of multiple equilibria and proposes the partial identification approach. Section 4 develops the subnetwork approach, where we derive the bounds from subnetworks and analyze their properties. Section 5 discusses the estimation methods. Section 6 discusses how to compute the bounds. Section 7 conducts a Monte Carlo study, and Section 8 concludes the paper.

2 A Model of Network Formation

In this section, we develop the network formation model. Let $[n] = \{1, 2, ..., n\}$ be the set of individuals who can form links. The links are undirected. This is a natural setting in the context of friendship networks, and for that reason we call linked individuals friends.

[2] See, e.g., Chernozhukov, Hong and Tamer (2007), Andrews and Soares (2010), Romano and Shaikh (2010), Andrews and Jia (2012) among others.

[3] Leung (2015) and Ridder and Sheng (2018) introduce independent privately known shocks so links formed across agents are conditional independent. Leung (2018) establishes a weak law for network moments under conditions that restrict strategic interaction effects to subnetworks of finite size. Menzel (2017) extends his idea on matching (Menzel, 2015) to network formation and provides conditions under which the distribution of link frequency has a limiting approximation where the dependence between links can be captured by certain sufficient statistics.

The links form a network, which we denote by $G \in \mathcal{G}$. It is an $n \times n$ binary matrix, where $G_{ij} = 1$ if i and j are friends, and 0 otherwise for all $i \neq j$. Since we consider undirected links, G is a symmetric matrix. We normalize $G_{ii} = 0$ for all i.

Utility Each individual i has a $d_x \times 1$ vector of observed attributes X_i (e.g., gender, age, race) and an $(n-1) \times 1$ vector of unobserved (to researchers) preferences $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{i,i-1}, \varepsilon_{i,i+1}, \ldots, \varepsilon_{in})'$, where ε_{ij} is i's preference for link ij. Let $X = (X_1', \ldots, X_n')'$ and $\varepsilon = (\varepsilon_1', \ldots, \varepsilon_n')'$. The utility of individual i in a network in general depends on the network configuration G, the observed attributes X, and i's unobserved preferences ε_i, i.e.,

$$U_i(G, X, \varepsilon_i).$$

For any $i \neq j$, we decompose G into (G_{ij}, G_{-ij}), where $G_{-ij} \in \mathcal{G}_{-ij}$ is the network obtained from G by removing link ij. Then the marginal utility of i from forming a link with j is $U_i(1, G_{-ij}, X, \varepsilon_i) - U_i(0, G_{-ij}, X, \varepsilon_i)$.

In this paper, we consider the utility specification

$$\begin{aligned} U_i(G, X, \varepsilon_i) &= \sum_{j=1}^{n} G_{ij} \left(u(X_i, X_j; \beta) + \varepsilon_{ij} \right) + \frac{1}{n-2} \sum_{j=1}^{n} \sum_{\substack{k=1 \\ k \neq i,j}}^{n} G_{ij} G_{jk} \gamma_1 \\ &\quad + \frac{1}{n-2} \sum_{j=1}^{n} \sum_{k=j+1}^{n} G_{ij} G_{ik} G_{jk} \gamma_2, \end{aligned} \tag{1}$$

where $u(X_i, X_j; \beta) = \beta_0 + \beta_1' X_i + \beta_2' |X_i - X_j|$. In this specification, the first term is the utility (net cost) from direct friends, where the term $|X_i - X_j|$ is to capture the homophily effect, which says that people tend to make friends with those who are similar to them (Currarini, Jackson and Pin (2009), Christakis et al. (2010)). In addition to the direct-friend effect, (1) also allows for the effect of indirect friends. The second term in (1) captures the utility from i's friends of friends, and the third term captures the additional utility if i and i's friend have friends in common,[4] where γ_1 and γ_2 are constants. Hence, if we consider the marginal utility of i from forming

[4]The latter is motivated by the clustering hypothesis, which says that if two individuals have friends in common, they are more likely to be friends than if links are formed randomly (Jackson and Rogers (2007), Jackson (2008), Christakis et al. (2010), Jackson et al. (2012)).

a link with j, which is given by

$$\Delta U_{ij}(G_{-ij}, X_i, X_j, \varepsilon_{ij}) = u(X_i, X_j; \beta) + \frac{1}{n-2} \sum_{\substack{k=1 \\ k \neq i,j}}^{n} G_{jk}\gamma_1 + \frac{1}{n-2} \sum_{\substack{k=1 \\ k \neq i,j}}^{n} G_{ik}G_{jk}\gamma_2 + \varepsilon_{ij}, \tag{2}$$

then it consists of not only the direct utility from j, but also the indirect utility from j's other friends and i, j's friends in common. This utility function follows closely the specification in Christakis et al. (2010).[5] It is also related to the specifications in Mele (2017) and Goyal and Joshi (2006), but is more general than both.[6]

Note that the effects of friends of friends and friends in common are normalized by $n-2$. Such a normalization reflects the fact that the marginal utility of i from a link with j should be influenced by a single other friend of j to a smaller extent if the network becomes larger. It is worth pointing out that the normalization does not result in diminishing network effects. The simulations in Section 7 provide evidence that even in large networks the specification in (2) yields link formation probabilities that differ substantially from those in a dyadic regression without strategic interactions (i.e., $\gamma_1 = \gamma_2 = 0$). The normalization ensures that the network effects do not grow to infinity and yield either a complete network or an empty network.

Equilibrium Given the utilities, individuals choose friends simultaneously as in a link-announcement game (Myerson (1991), Jackson (2008)). We assume that individuals observe X and ε, so it is a complete-information game. Depending on whether transfers are allowed for, each individual announces a set of *intended links* or *intended transfers*. Under nontransferable utility (NTU), a link is formed if both individuals intend to form it, while under transferable utility (TU) a link is formed if the sum of the two transfers for it is nonnegative.

The equilibrium concept we consider in this paper is *pairwise stability* (Jackson and Wolinsky (1996) for NTU, Bloch and Jackson (2006, 2007) for TU). We say that a network is pairwise stable if no pair of individuals wants to create a new link, and no individual wants to sever an existing link. Formally,

[5] Christakis et al. (2010) allow for nonlinear effects from friends of friends and friends in common. Our specification is a linear version of theirs. However, with linearity we can establish the existence of equilibrium, which is an open question for the specification they use.

[6] Mele (2011) considers a linear utility function which does not allow for the effects of friends in common. Goyal and Joshi (2006) assume that the direct-friend effects do not depend on individual attributes.

Definition 2.1 *A network G is pairwise stable (PS) under NTU if*

1. *for any $G_{ij} = 1$, $\Delta U_{ij}(G_{-ij}, X_i, X_j, \varepsilon_{ij}) \geq 0$ and $\Delta U_{ji}(G_{-ji}, X_j, X_i, \varepsilon_{ji}) \geq 0$;*
2. *for any $G_{ij} = 0$, $\Delta U_{ij}(G_{-ij}, X_i, X_j, \varepsilon_{ij}) > 0 \implies \Delta U_{ji}(G_{-ji}, X_j, X_i, \varepsilon_{ji}) < 0$.*

Definition 2.2 *A network G is pairwise stable (PS) under TU if*

1. *for any $G_{ij} = 1$, $\Delta U_{ij}(G_{-ij}, X_i, X_j, \varepsilon_{ij}) + \Delta U_{ji}(G_{-ji}, X_j, X_i, \varepsilon_{ji}) \geq 0$;*
2. *for any $G_{ij} = 0$, $\Delta U_{ij}(G_{-ij}, X_i, X_j, \varepsilon_{ij}) + \Delta U_{ji}(G_{-ji}, X_j, X_i, \varepsilon_{ji}) \leq 0$.*

In the sequel we consider both NTU and TU and use the term "pairwise stability" to mean pairwise stability under NTU or TU, depending on the context.

Since we allow for utility interdependence, the pairwise stability condition leads to a simultaneous discrete choice model, i.e.,

$$G_{ij} = 1\left\{\Delta U_{ij}(G_{-ij}, X_i, X_j, \varepsilon_{ij}) \geq 0, \Delta U_{ji}(G_{-ji}, X_j, X_i, \varepsilon_{ji}) \geq 0\right\}, \ \forall i \neq j, \tag{3}$$

under NTU and

$$G_{ij} = 1\left\{\Delta U_{ij}(G_{-ij}, X_i, X_j, \varepsilon_{ij}) + \Delta U_{ji}(G_{-ji}, X_j, X_i, \varepsilon_{ji}) \geq 0\right\}, \ \forall i \neq j, \tag{4}$$

under TU,[7] where the choice of a link G_{ij} depends on the choices of others G_{-ij}. This indicates that we cannot treat each link as a single observation and use a dyadic regression because G_{-ij} is endogenous in the model, and therefore can be correlated with $(\varepsilon_{ij}, \varepsilon_{ji})$. What further complicates the statistical inference of (3) and (4) is that there may be multiple equilibria, which will affect the identification of the parameters.

The existence of pairwise stable networks is also not guaranteed. According to Jackson and Watts (2002, Lemma 1), for any utility function there is either a PS network or a closed cycle.[8,9] In the appendix we give an example where there is no

[7] Equations (3) and (4) differ slightly from Definitions 2.1 and 2.2 in the indifference case, but the discrepency is negligible when ε follows a continuous distribution.

[8] A *closed cycle* is a collection of two or more distinct networks such that: (i) for any two networks in the collection there is an improving path from one to the other; and (ii) no improving path starting from a network in the collection leads to a network outside. Here an *improving path* is a sequence of networks in which two consecutive networks differ by one link, and adding (or deleting) the link in the successive network is beneficial for the individuals involved. See Jackson and Watts (2002) for rigorous definitions.

[9] The original result in Jackson and Watts (2002) was proved under NTU. It is easy to show that their result also holds under TU.

PS network, but a closed cycle. A closed cycle represents a situation in which *for the given utilities* individuals never reach a stable state and constantly switch between forming and severing links, which is unlikely to occur in real applications. To ensure that our model yields an appropriate solution, we need a utility function such that for any parameter value, X and ε, there exists a PS network.

Most results in the network literature on the existence of PS networks do not allow for heterogeneity among individuals and thus are unsuitable for our analysis.[10] Jackson and Watts (2001) and Hellmann (2012) provide general conditions under which a PS network exists. We apply their conditions and provide existence results for the utility function in (1). The insight of these results is that (1) under TU the model permits a representation as a potential game (Monderer and Shapley, 1994), and (2) under NTU, with the additional assumption that links are strategic complements, the model is a supermodular game (Milgrom and Roberts, 1990), so the existence of equilibrium follows from the fixed-point theorem for isotone mappings (Topkis, 1979). Detailed proofs are given in the appendix.

Proposition 2.1 *Suppose that the utility function is as in (1). Under TU, for any function u and any constants γ_1 and γ_2 in $\mathbb{R}$, there is no closed cycle, so a PS network must exist.*

Proposition 2.2 *Suppose that the utility function is as in (1). Under NTU, for any function u and any constants $\gamma_1 \geq 0$ and $\gamma_2 \geq 0$, there is no closed cycle, so a PS network must exist.*

Remark 2.1 *The existence results in Propositions 2.1-2.2 can be extended to generalizations of the utility specification in (1) where γ_1 and γ_2 depend on the attributes. Suppose that the coefficients of the terms $G_{ij}G_{jk}$ and $G_{ij}G_{ik}G_{jk}$ in (1) take the form of $\gamma_1(X_i, X_j, X_k)$ and $\gamma_2(X_i, X_j, X_k)$. If $\gamma_1(X_i, X_j, X_k)$ is symmetric in X_i and X_k, and $\gamma_2(X_i, X_j, X_k)$ is symmetric in X_i, X_j, and X_k, the result in Proposition 2.1 still holds. Moreover, the result in Proposition 2.2 holds if $\gamma_1(X_i, X_j, X_k) \geq 0$ and $\gamma_2(X_i, X_j, X_k) \geq 0$ for all X_i, X_j, and X_k.*

[10] See, for example, Belleflamme and Bloch (2004), Goyal and Joshi (2006).

3 Partial Identification

In this section, we examine the general framework that we use to identify the model. After introducing the data generating process, we discuss multiple equilibria, the main problem in identification. Then we show how much we can learn about the parameters without imposing any restrictions on the equilibrium selection.

We consider the following data generating process. Let n be an integer generated from a distribution on $\{2, 3, \ldots\}$. We draw n individuals at random from a super-population. Each individual i is associated with a vector of observed attributes $X_{n,i}$ and a vector of unobserved preferences $\varepsilon_{n,i}$. We let these n individuals form links, and a PS network $G_n = (G_{n,ij})_{i \neq j}$ emerges. For notational convenience, we define $X_{n,ij} = (X_{n,i}, X_{n,j})$ to be the attributes of a pair (i, j) and $X_n = (X_{n,ij})_{i \neq j}$ the attribute profile of all the pairs. We observe T independent networks and their attribute profiles (G_{n_t}, X_{n_t}, n_t), $t = 1, \ldots, T$.

Throughout the paper we make the following assumptions.

Assumption 1 (Data generating process) *(i) We have an i.i.d. sample of* (G_{n_t}, X_{n_t}, n_t), $t = 1, \ldots, T$. *Assume* $T \to \infty$. *(ii)* X_{n_t} *and* ε_{n_t} *are independent for all* $t = 1, \ldots, T$. *(iii)* $\varepsilon_{n_t,ij}$ *for all* $i \neq j$ *and* $t = 1, \ldots, T$ *are i.i.d. from a distribution with CDF* $F(\varepsilon_{ij}; \theta_\varepsilon)$ *supported on* $\mathbb{R}$ *that is absolutely continuous with respect to the Lebesgue measure.* $F(\varepsilon_{ij}; \theta_\varepsilon)$ *is continuously differentiable in the finite-dimensional parameter* $\theta_\varepsilon \in \Theta_\varepsilon$.

Assumption 2 (Utility) *The marginal utility of* i *from forming a link with* j *has a form* $\Delta U_{ij}(G_{n,-ij}, X_{n,ij}, \varepsilon_{n,ij}; \theta_u)$ *as specified in (2), where* $\theta_u = (\beta, \gamma) \in \Theta_u$ *denotes the utility function parameter.*

The parameter of interest is $\theta = (\theta_u, \theta_\varepsilon) \in \Theta_u \times \Theta_\varepsilon = \Theta$.

For a given attribute profile X_n and preference profile ε_n, the model yields a collection of PS networks, denoted by $\mathcal{PS}(\Delta U_n(X_n, \varepsilon_n))$, where $\Delta U_n(X_n, \varepsilon_n) = \{\Delta U_{ij}(G_{n,-ij}, X_{n,ij}, \varepsilon_{n,ij}), \forall G_{n,-ij} \in \mathcal{G}_{n,-ij}, \forall i \neq j\}$ represents the marginal-utility profile, and $\mathcal{G}_{n,-ij}$ denotes the set of all possible $G_{n,-ij}$. To complete the model, suppose there is an equilibrium selection mechanism that selects a network from the collection of PS networks. Let $\lambda_n(g_n | \mathcal{PS}(\Delta U_n(X_n, \varepsilon_n)), X_n, \varepsilon_n)$ be the probability with which a network g_n is selected from the PS collection $\mathcal{PS}(\Delta U_n(X_n, \varepsilon_n))$. Then

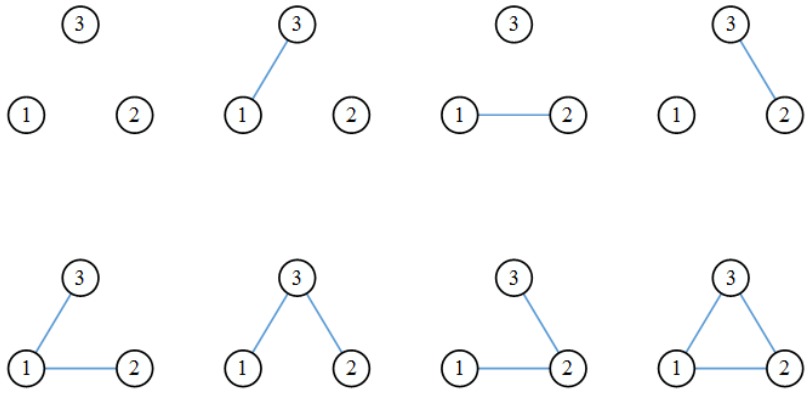

Figure 1: Networks of Three Individuals

conditional on X_n the probability of observing the network g_n is

$$\Pr\left(G_n = g_n | \, X_n\right) = \int \lambda_n \left(g_n | \mathcal{PS}\left(\Delta U_n\left(X_n, \varepsilon_n\right)\right), X_n, \varepsilon_n\right) dF\left(\varepsilon_n\right) \tag{5}$$

Equation (5) is similar to what Ciliberto and Tamer (2009) establish in entry games and Bajari, Hong, and Ryan (2010) in discrete games with complete information.

For the network formation game described in Section 2, the presence of multiple equilibria is prevalent because of the interdependence of marginal utilities across links.[11] We illustrate multiple equilibria in Example 3.1.

Example 3.1 *Consider networks of size $n = 3$. Figure 1 shows the 8 possible network configurations. Consider the utility function as in (1) with $u\left(X_i, X_j; \beta\right) = u\left(X_j, X_i; \beta\right)$, $\gamma_1 > 0$, and $\gamma_2 > 0$. Abbreviate $u\left(X_i, X_j; \beta\right)$ as u_{ij}. For simplicity we assume $\varepsilon_{ij} = \varepsilon_{ji}$, so $\varepsilon = \left(\varepsilon_{12}, \varepsilon_{23}, \varepsilon_{13}\right) \in \mathbb{R}^3$. Given the utility specification, we calculate all possible collections of PS networks under TU. The regions of ε that correspond to each collection of PS networks are presented in Figure 2, where a network g is represented by the vector $\left(g_{12}, g_{23}, g_{13}\right) \in \{0,1\}^3$. In this example, all the 8 networks belong to certain collections of multiple PS networks.*

If all network outcomes belong to certain collections of multiple equilibria, as illustrated in Example 3.1, without knowledge of the equilibrium selection no network outcome can yield moment restrictions that point identify the true θ_0. In this paper, we leave the equilibrium selection completely unrestricted and propose a partial iden-

[11] Note that if there is no utility interdependence, i.e., $\Delta U_{ij}\left(G_{n,-ij}, X_{n,ij}, \varepsilon_{n,ij}\right) = \Delta U_{ij}\left(X_{n,ij}, \varepsilon_{n,ij}\right)$, then a pairwise stable network must be unique.

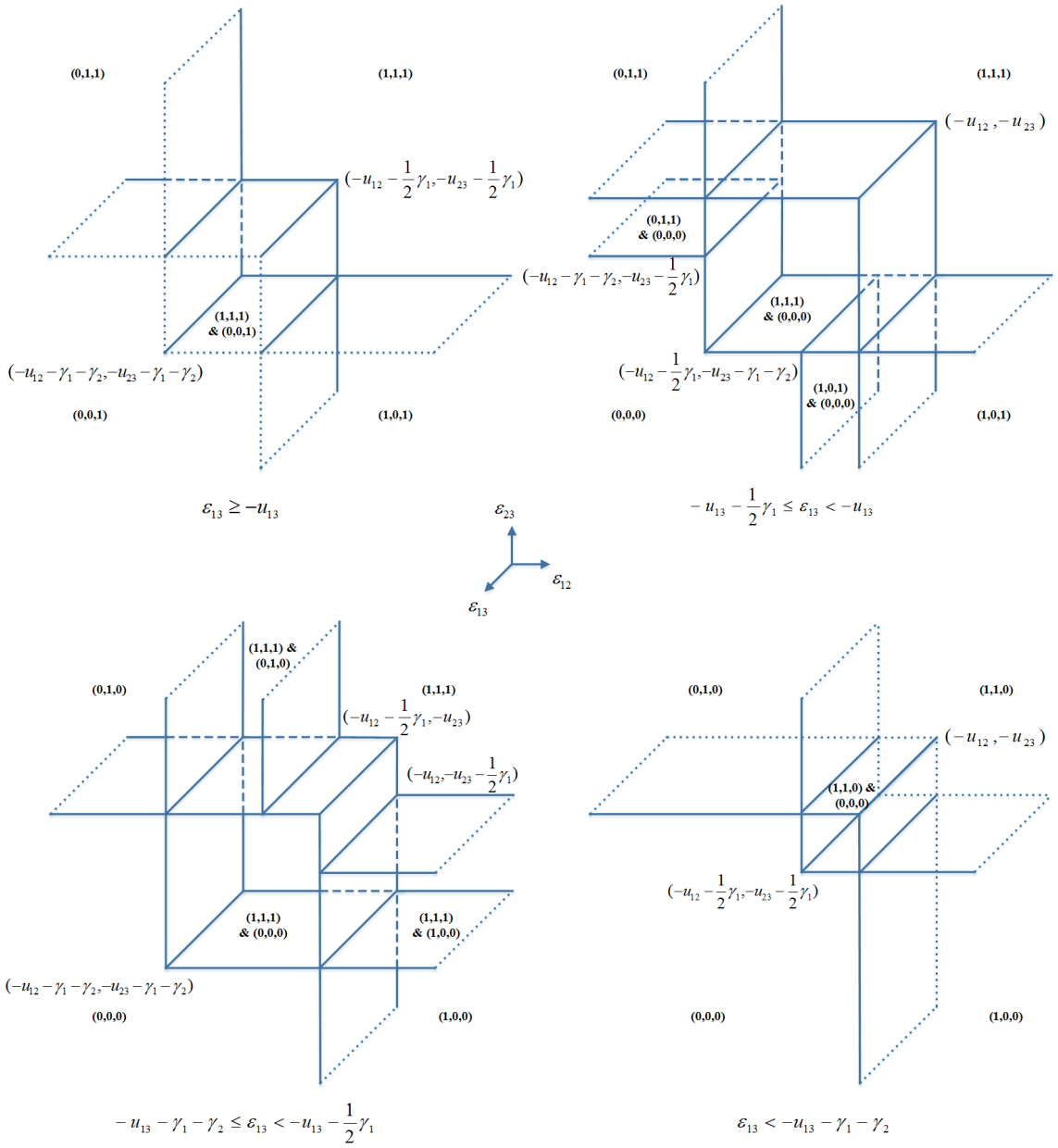

Figure 2: All Possible Equilibria and the Partition of the ε Space

tification approach.[12] Following closely Ciliberto and Tamer (2009), we divide the integral in (5) into two parts, depending on whether there is a unique equilibrium or multiple equilibria,

$$\begin{aligned}\Pr\left(G_n = g_n \middle| X_n\right) = & \int_{g_n \in \mathcal{PS}(\Delta U_n(X_n,\varepsilon_n)) \& |\mathcal{PS}(\Delta U_n(X_n,\varepsilon_n))|=1} dF\left(\varepsilon_n\right) \\ & + \int_{g_n \in \mathcal{PS}(\Delta U_n(X_n,\varepsilon_n)) \& |\mathcal{PS}(\Delta U_n(X_n,\varepsilon_n))|\geq 2} \lambda_n\left(g_n | \mathcal{PS}\left(\Delta U_n\left(X_n,\varepsilon_n\right)\right), X_n, \varepsilon_n\right) dF\left(\varepsilon_n\right),\end{aligned} \tag{6}$$

Note that the selection probability is equal to 1 if a network is a unique equilibrium. If there are multiple equilibria, the selection probability, though unknown, lies between 0 and 1. Replacing the selection probability with these bounds, we derive an upper and lower bound for $\Pr\left(G_n = g_n \middle| X_n\right)$, i.e.,

$$\Pr\left(G_n = g_n \middle| X_n\right) \leq \int_{g_n \in \mathcal{PS}(\Delta U_n(X_n,\varepsilon_n))} dF\left(\varepsilon_n\right), \tag{7}$$

and

$$\Pr\left(G_n = g_n \middle| X_n\right) \geq \int_{g_n \in \mathcal{PS}(\Delta U_n(X_n,\varepsilon_n)) \& |\mathcal{PS}(\Delta U_n(X_n,\varepsilon_n))|=1} dF\left(\varepsilon_n\right). \tag{8}$$

The upper bound is the probability that network g_n is PS, and the lower bound is the probability that network g_n is uniquely PS. These are the best possible bounds for $\Pr\left(G_n = g_n \middle| X_n\right)$ because the selection probability in (6) can be any value between 0 and 1.

Unfortunately, the computation of these bounds suffers from the curse of dimensionality in large networks. In particular, the lower bound in (8) is computationally infeasible if n is large. This is because to compute the lower bound, we need to check pairwise stability for $2^{n(n-1)/2}$ possible networks.[13] This is computationally in-

[12] An alternative approach is to make additional assumptions about the equilibrium selection. In network formation, one way to do this is to consider a sequential model as in Jackson and Watts (2002) (see also Christakis et al. (2010) and Mele (2011)), which typically yields a unique network or a unique stationary distribution over networks, depending on the setting. Alternatively, one can assume a more general equilibrium selection mechanism, for example, by specifying a parametric form (Bajari, Hong, and Ryan (2010)) or considering a nonparametric equilibrium selection (Bajari, Hahn, Hong, and Ridder (2011)). In the game we consider a fully nonparametric equilibrium selection is not identified. Certain restrictions must be imposed on it in order to achieve identification..

[13] Unlike the upper bound, the lower bound has no closed form and needs to be computed by simulation. For each simulated ε_n, we need to check whether a network is uniquely pairwise stable, which amounts to checking pairwise stability for all possible networks.

tractable even for a moderate value of n. For example, in the case of 20 people, the number of possible networks is $2^{190} \approx 10^{57}$.

4 Partial Identification from Subnetworks

4.1 Bounds from Subnetworks

We propose a novel approach to reduce the dimensionality of the problem. The idea is to derive bounds for certain parts of a network, called subnetworks. A subnetwork is the restriction of a network to a subset of the individuals. To be precise, let G_n be a network of n nodes. For any subset $A \subseteq [n]$, we say that $G_{n,A}$ is the *subnetwork* of G_n in A if it consists of the edges in G_n that connect two nodes in A, i.e., $G_{n,A} = (G_{n,ij})_{i,j \in A, i \neq j}$. Moreover, we define $G_{n,-A}$ to be the *complement* of $G_{n,A}$, i.e., the remainder of G_n after the edges in $G_{n,A}$ are deleted, so $G_{n,-A} = (G_{n,ij})_{i \notin A \cup j \notin A, i \neq j}$. In matrix notation, the subnetwork $G_{n,A}$ corresponds to the submatrix of G_n with rows and columns in A, and its complement $G_{n,-A}$ is the remainder of G_n after the submatrix in A is deleted. The sets of all possible $G_{n,A}$ and $G_{n,-A}$ are denoted by $\mathcal{G}_{n,A}$ and $\mathcal{G}_{n,-A}$.

For any fixed subset $A \subseteq [n]$, it is clear from the decomposition $G_n = (G_{n,A}, G_{n,-A})$ that the distribution of the subnetwork $G_{n,A}$ is simply a marginal distribution of the network G_n. That is, conditional on X_n the probability of observing a subnetwork $g_{n,A}$ is

$$\begin{aligned} \Pr\left(G_{n,A} = g_{n,A} | \, X_n\right) &= \sum_{g_{n,-A}} \Pr\left(G_{n,A} = g_{n,A}, G_{n,-A} = g_{n,-A} | \, X_n\right) \\ &= \int \sum_{g_{n,-A}} \lambda_n \left(g_{n,A}, g_{n,-A} | \, \mathcal{PS}\left(\Delta U_n\left(X_n, \varepsilon_n\right)\right), X_n, \varepsilon_n\right) dF\left(\varepsilon_n\right). \end{aligned} \tag{9}$$

The summation over λ_n in (9) is unknown unless all the PS networks in the set $\mathcal{PS}\left(\Delta U_n\left(X_n, \varepsilon_n\right)\right)$ have the same subnetwork $g_{n,A}$. Define

$$\mathcal{PS}_A\left(\Delta U_n\left(X_n, \varepsilon_n\right)\right) = \{g_{n,A} \in \mathcal{G}_{n,A} : \exists g_{n,-A} \in \mathcal{G}_{n,-A}, (g_{n,A}, g_{n,-A}) \in \mathcal{PS}(\Delta U_n(X_n, \varepsilon_n))\}$$

to be the set of subnetworks in A that can be part of a PS network in the set

$\mathcal{PS}(\Delta U_n(X_n, \varepsilon_n))$. Following the idea in the previous section, we can derive an upper and lower bound for the subnetwork choice probability $\Pr(G_{n,A} = g_{n,A} | X_n)$.

Proposition 4.1 *For any subset $A \subseteq [n]$ and any subnetwork $g_{n,A}$, the subnetwork choice probability $\Pr(G_{n,A} = g_{n,A} | X_n)$ is bounded by*

$$H_{2n}(g_{n,A}, X_n) \leq \Pr(G_{n,A} = g_{n,A} | X_n) \leq H_{1n}(g_{n,A}, X_n) \tag{10}$$

where

$$\begin{aligned} H_{1n}(g_{n,A}, X_n) &= \int_{g_{n,A} \in \mathcal{PS}_A(\Delta U_n(X_n, \varepsilon_n))} dF(\varepsilon_n) \\ H_{2n}(g_{n,A}, X_n) &= \int_{g_{n,A} \in \mathcal{PS}_A(\Delta U_n(X_n, \varepsilon_n)) \& |\mathcal{PS}_A(\Delta U_n(X_n, \varepsilon_n))| = 1} dF(\varepsilon_n) \end{aligned}$$

Proof. Divide the integral in (9) into two parts, depending on whether the PS networks have a unique subnetwork or multiple subnetworks,

$$\begin{aligned} &\Pr(G_{n,A} = g_{n,A} | X_n) \\ &= \int_{\substack{g_{n,A} \in \mathcal{PS}_A(\Delta U_n(X_n, \varepsilon_n)) \\ \& |\mathcal{PS}_A(\Delta U_n(X_n, \varepsilon_n))| = 1}} dF(\varepsilon_n) \\ &\quad + \int_{\substack{g_{n,A} \in \mathcal{PS}_A(\Delta U_n(X_n, \varepsilon_n)) \\ \& |\mathcal{PS}_A(\Delta U_n(X_n, \varepsilon_n))| \geq 2}} \sum_{g_{n,-A}} \lambda_n(g_{n,A}, g_{n,-A} | \mathcal{PS}(\Delta U_n(X_n, \varepsilon_n)), X_n, \varepsilon_n) \, dF(\varepsilon_n), \end{aligned} \tag{11}$$

Replacing the sum term in (11) by 0 and 1 yields (10). ■

When we observe a subnetwork $g_{n,A}$, it must be part of a PS network for some complement $g_{n,-A}$ and such a PS network $(g_{n,A}, g_{n,-A})$ must be selected by the equilibrium selection mechanism. Without information about the equilibrium selection, the probability of observing the subnetwork $g_{n,A}$ is naturally bounded above by the probability that the subnetwork $g_{n,A}$ can be part of a PS network for some complement $g_{n,-A}$, which gives the upper bound in (10), and bounded below by the probability that only the subnetwork $g_{n,A}$ can be part of a PS network for some complement $g_{n,-A}$, which gives the lower bound in (10). If the selection mechanism *always* selects a PS network with the subnetwork $g_{n,A}$ whenever such a PS network is available, the subnetwork choice probability $\Pr(G_{n,A} = g_{n,A} | X_n)$ is equal to the upper bound. On the

other hand, if the selection mechanism *never* selects a PS network with the subnetwork $g_{n,A}$ unless all the PS networks have the same subnetwork $g_{n,A}$, the subnetwork choice probability $\Pr\left(G_{n,A} = g_{n,A} | \, X_n\right)$ is equal to the lower bound. This implies that the bounds in (10) provide the best possible bounds for $\Pr\left(G_{n,A} = g_{n,A} | \, X_n\right)$.

In contrast to the lower bound in (8), these subnetwork bounds can be computed even in large networks as long as the subnetworks are chosen to be small. Details of the computation are discussed in Section 6.

Example 4.1 (Example 3.1 continued) *Assume the same setting as in Example 3.1. We calculate the upper and lower bounds in (10) for the subnetwork* $G_{12} = 1$ *(suppress the subscript* n*). For this pair subnetwork, its complement* g_{-12} *takes four possible values* $\mathcal{G}_{-12} = \{(1,1), (1,0), (0,1), (0,0)\}$*. The upper bound for* $\Pr\left(G_{12} = 1 | \, X\right)$ *is obtained from the regions of* ε *where the subnetwork* $G_{12} = 1$ *and some complement* $g_{-12} \in \mathcal{G}_{-12}$ *form a PS network. From Figure 2 we can see that there are 10 such regions. Define* $R_1(X)$ *to be the union of these regions, i.e.,*

$$
\begin{aligned}
R_1(X) &= \left\{\varepsilon \in \mathbb{R}^3 : \exists g_{-12} \in \mathcal{G}_{-12},\ (1, g_{-12}) \text{ is PS} \right\} \\
&= \Big\{\varepsilon \in \mathbb{R}^3 : (1,1,1) \text{ uniquely PS, or } (1,1,0) \text{ uniquely PS, or} \\
&\qquad (1,0,1) \text{ uniquely PS, or } (1,0,0) \text{ uniquely PS, or} \\
&\qquad \{(1,1,1),(1,0,0)\} \text{ PS, or } \{(1,1,1),(0,1,0)\} \text{ PS, or} \\
&\qquad \{(1,1,1),(0,0,1)\} \text{ PS, or } \{(1,1,1),(0,0,0)\} \text{ PS, or} \\
&\qquad \{(1,1,0),(0,0,0)\} \text{ PS, or } \{(1,0,1),(0,0,0)\} \text{ PS} \qquad \Big\}
\end{aligned}
$$

The integral over the region $R_1(X)$ *gives the upper bound for* $\Pr\left(G_{12} = 1 | \, X\right)$,

$$
\Pr\left(G_{12} = 1 | \, X\right) \leq \int_{R_1(X)} dF(\varepsilon).
$$

As for the lower bound, it is obtained from the regions included in the upper bound where only the subnetwork $G_{12} = 1$ *with some complement* $g_{-12} \in \mathcal{G}_{-12}$ *can form a PS network, or equivalently, the subnetwork* $G_{12} = 0$ *with any complement* $g_{-12} \in \mathcal{G}_{-12}$ *cannot form a PS network. From Figure 2 there are* 5 *such regions. Define* $R_2(X)$ *to be the union of these* 5 *regions, i.e.,*

$$
R_2(X) = \left\{\varepsilon \in \mathbb{R}^3 : \exists g_{-12} \in \mathcal{G}_{-12},\ (1, g_{-12}) \text{ is PS, and} \right.
$$

$$
\begin{aligned}
& \forall g_{-12} \in \mathcal{G}_{-12},\ (0, g_{-12}) \ \textit{is not PS}\ \} \\
= \ & \{\varepsilon \in \mathbb{R}^3 : (1,1,1) \ \textit{uniquely PS, or}\ (1,1,0) \ \textit{uniquely PS, or} \\
& (1,0,1) \ \textit{uniquely PS, or}\ (1,0,0) \ \textit{uniquely PS, or} \\
& \{(1,1,1),(1,0,0)\} \ \textit{PS} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \}
\end{aligned}
$$

The integral over the region $R_2(X)$ *gives the lower bound for* $\Pr(G_{12} = 1|\, X)$,

$$
\Pr(G_{12} = 1|\, X) \geq \int_{R_2(X)} dF(\varepsilon).
$$

When we use the subnetwork inequalities in (10) in the estimation, we need to make sure that the estimates do not depend on how we label the individuals. Note that the utility specification in (1) and the pairwise stability conditions in (3)-(4) do not depend on the labels. We can get around any labeling issue by imposing an additional mild assumption that the equilibrium selection mechanism is invariant to permutations of the labels. To see this, consider a network G_n together with its attribute profile X_n and call (G_n, X_n) an attributed network or simply network. Let $(\mathcal{G}_n, \mathcal{X}_n)$ be the set of all possible (G_n, X_n). We say that an equilibrium selection mechanism λ_n is invariant to permutations of the labels if for any permutation π over $[n]$, any network $(g_n, x_n) \in (\mathcal{G}_n, \mathcal{X}_n)$, and any preference profile ε_n, we have

$$
\lambda_n(g_n^\pi|\, \mathcal{PS}(\Delta U_n(x_n^\pi, \varepsilon_n^\pi)), x_n^\pi, \varepsilon_n^\pi) = \lambda_n(g_n|\, \mathcal{PS}(\Delta U_n(x_n, \varepsilon_n)), x_n, \varepsilon_n), \tag{12}
$$

where $(g_n^\pi, x_n^\pi) \in (\mathcal{G}_n, \mathcal{X}_n)$ is the network obtained from (g_n, x_n) by π, i.e., the network obtained by permuting the rows and columns of (g_n, x_n) according to π,[14] and ε_n^π is the preference profile obtained from ε_n by π. The condition in (12) means that if a set of multiple equilibria (in terms of (g_n, x_n)) can be obtained from another by relabeling the individuals, the selection probabilities should remain the same under the relabeling.[15] For a given network (G_n, X_n), if for any permutation π over $[n]$ the

[14] Specifically, the (i,j) element of (g_n^π, x_n^π) is equal to the $(\pi(i), \pi(j))$ element of (g_n, x_n) for all $i \neq j$.

[15] For instance, in Example 3.1, there are two regions of ε where in one region $(1,0,0)$ and $(1,1,1)$ are the PS networks and in the other $(0,1,0)$ and $(1,1,1)$ are the PS networks (both are in the lower left of Figure 2). In this example $X_1 = X_3$, so the latter collection of multiple equilibria can be obtained from the former by swapping the labels of players 1 and 3. The condition in (12) then means that the probability of selecting $(1,0,0)$ from the collection $\{(1,0,0),(1,1,1)\}$ should be equal to the probability of selecting $(0,1,0)$ from the collection $\{(0,1,0),(1,1,1)\}$.

network (G_n^π, X_n^π) obtained from (G_n, X_n) by π has the same distribution as (G_n, X_n), i.e., $(G_n^\pi, X_n^\pi) \stackrel{d}{=} (G_n, X_n)$, the network (G_n, X_n) is exchangeable (Kallenberg, 2005). It is clear that any network (G_n, X_n) generated under the utility specification in (1) and an equilibrium selection mechanism λ_n that satisfies the condition (12) is exchangeable.

The exchangeability of (G_n, X_n) implies that any two subnetworks $(G_{n,A}, X_{n,A})$ and $(G_{n,A'}, X_{n,A'})$ of the same size (i.e., $|A| = |A'|$) have the same distribution, so it suffices to consider the subnetwork in $A = [a]$, denoted by $(G_{n,a}, X_{n,a})$, and its choice probability $\Pr(G_{n,a} = g_a | X_{n,a} = x_a, X_{n,-a})$. Moreover, if a subnetwork (g_a, x_a) can be obtained from another subnetwork (g'_a, x'_a) by a permutation over $[a]$ (in this case we say that (g_a, x_a) is *isomorphic* to (g'_a, x'_a)), by exchangeability the choice probabilities evaluated at such subnetworks are equal. Therefore, we can view $\Pr(G_{n,a} = g_a | X_{n,a} = x_a, X_{n,-a})$ as a function of unlabeled subnetworks (g_a, x_a) and the "=" symbol in the choice probability as an isomorphism.[16] This ensures that the choice probabilities we estimate do not depend on the labels. Note that considering unlabeled subnetworks does not invalidate the subnetwork inequalities in (10) because under Assumptions 1-2 the bounds in (10) evaluated at isomorphic subnetworks are also equal. Denote the bounds for an unlabeled subnetwork (g_a, x_a) by $H_{1n}(g_a, x_a, X_{n,-a})$ and $H_{2n}(g_a, x_a, X_{n,-a})$.

A natural question about the subnetwork bounds is whether the bounds from a subnetwork remain informative when the network becomes large. In the rest of the section, we analyze the properties of these subnetwork bounds when we vary the size of a network. Under the utility specification in (1), we establish in Proposition 4.2 that for any subnetwork its upper and lower bounds are strictly bounded away from 0 and 1 for all network sizes.

Proposition 4.2 *Suppose that Assumptions 1-2 are satisfied. For any $a \leq n$ and any $(g_a, x_a) \in (\mathcal{G}_a, \mathcal{X}_a)$, the bounds $H_{1n}(g_a, x_a, X_{n,-a})$ and $H_{2n}(g_a, x_a, X_{n,-a})$ in (10) satisfy*

$$\bar{H}_2(g_a, x_a) \leq H_{2n}(g_a, x_a, X_{n,-a}) \leq H_{1n}(g_a, x_a, X_{n,-a}) \leq \bar{H}_1(g_a, x_a)$$

[16] The use of isomorphism and unlabeled subnetworks is needed only for discrete X. If X is continuous, two individuals have distinct X with probability 1, so no subnetworks are isomorphic and labeled and unlabeled subnetworks are the same.

for some deterministic functions $\bar{H}_1(g_a, x_a)$ *and* $\bar{H}_2(g_a, x_a)$ *such that* $0 < \bar{H}_2(g_a, x_a) < \bar{H}_1(g_a, x_a) < 1$.

Proof. See the appendix. ■

The boundedness property results from the fact that in the marginal utility specification in (2) the network effects from friends of friends and friends in common are bounded for all network sizes. For expositional convenience, we say that a subnetwork (respectively, complement) is PS for a complement (respectively, subnetwork) if each potential link in the subnetwork (respectively, complement) satisfies the stability condition in (3) or (4). If a subnetwork is PS for some PS complement, it must be PS under the most favorable network effects (i.e., the largest network effect for links being 1 and the smallest network effect for links being 0). Thus the upper bound H_{1n} is bounded above by the probability of the latter event, defined as the upper limit $\bar{H}_1$, which is strictly less than 1 because ε has a full support on $\mathbb{R}$. On the other hand, if a subnetwork is PS under the least favorable network effects (i.e., the opposite of the most favorable case), it must be the unique PS subnetwork for all PS complements, so the lower bound H_{2n} is bounded below by the probability of the former event, defined as the lower limit $\bar{H}_2$, which is strictly above 0. These results ensure that the upper and lower bounds in (10) are nontrivial for all network sizes.

Remark 4.1 *The results in Proposition 4.2 are not restricted to the utility specification in (1). They hold for more general specifications, for example, where* γ_1 *and* γ_2 *depend on the attributes* X*, so long as the network effects from friends of friends and friends in common are bounded and the existence of a PS network is guaranteed as discussed in Remark 2.1.*

4.2 Identified Set

The inequalities in (10) are satisfied by the true parameter value θ_0, i.e.,

$$H_{2n}(g_a, x_a, X_{n,-a}; \theta_0) \leq \Pr(G_{n,a} = g_a | X_{n,a} = x_a, X_{n,-a}) \leq H_{1n}(g_a, x_a, X_{n,-a}; \theta_0) \tag{13}$$

for all (g_a, x_a), all $X_{n,-a}$ and all $a \leq n$. Let $\Theta_I(a)$ be the set of parameter values that satisfy the inequalities in (13) for subnetworks of size a,

$$\Theta_I(a) = \{\theta \in \Theta : (13) \text{ holds for the given } a \text{ with } \theta \text{ in place of } \theta_0\}. \tag{14}$$

We define the identified set as $\Theta_I = \bigcap_{a=2}^{\bar{a}} \Theta_I(a)$ for some positive integer $\bar{a}$.

The subnetwork bounds in (13) provide nontrivial bounds for the probability of observing a subnetwork. These bounds take into account not only the pairwise stability of the subnetwork, but also the pairwise stability of its complement. To obtain the bounds, we need to go back to the regions in the space of ε_n that correspond to each possible collection of PS networks, the same as needed for the full network bounds, and group the regions according to whether the subnetwork is PS (or uniquely PS), as illustrated in Example 4.1. All such regions of ε_n have boundaries that depend on θ (as seen in Figure 2), so the subnetwork bounds obtained by taking the integral over the relevant regions would generally vary with θ.

We obtain the set $\Theta_I(a)$ by comparing the subnetwork bounds evaluated at each value of θ with the observed probabilities of subnetworks. If a value of θ generates an upper bound that falls below the observed probability of a subnetwork, or a lower bound that exceeds the observed probability of a subnetwork, this value is excluded from the set $\Theta_I(a)$. Figure 2 and Example 4.1 indicate that extreme values of θ tend to shift both the upper and lower bounds to extreme values which are unlikely to cover the observed probability of a subnetwork, and thus would be excluded from the set $\Theta_I(a)$.[17] This implies that the set $\Theta_I(a)$ is a proper subset of Θ; the subnetwork bounds are informative about θ.

Compared with the bounds from a full network, the bounds from subnetworks of size $a < n$ are generally less informative about θ. The full network bounds provide additional inequality restrictions on θ by exploiting information from the joint distribution of a subnetwork and its complement. It is possible that certain values of θ generate full network bounds that cannot cover the observed joint probability of a subnetwork and its complement, but after aggregating over the complements they yield subnetwork bounds that cover the observed probability of the subnetwork. In general, the identified set Θ_I obtained from subnetworks up to size $\bar{a}$ provides a superset of the sharp identified set obtained from the full network.[18] By considering this superset we get around the computationally prohibitive full network bounds that

[17] The simulations in Section 7 provide similar evidence. Figure 3 shows that for $n \geq 25$ both the upper and lower bounds for pair subnetworks approach 1 as γ_2 increases. Unless the observed link probability is equal to 1, such extreme values of γ_2 would be excluded from the identified set.

[18] The sharp identified set is defined as the collection of θ under which there is an equilibrium selection mechanism that generates the observed distribution of full networks (Beresteanu, Molchanov and Molinari (2011)).

are required by the sharp set.

In practice, if n is large we may need to choose a small $\bar{a}$. In the simulation study in Section 7, we evaluate the performance of bounds from small subnetworks (e.g. $\bar{a} = 2$ or 3). We provide evidence that bounds from such small subnetworks are informative about θ, even in large networks.

5 Estimation

In this section, we discuss the estimation of the identified set Θ_I. This set is defined by the conditional moment inequalities

$$\begin{aligned} \mathbb{E}\left[1\left\{G_{n,a} = g_a\right\} - H_{1n}\left(g_a, X_{n,a}, X_{n,-a}; \theta\right) \middle| X_{n,a}, X_{n,-a}\right] &\leq 0 \\ \mathbb{E}\left[H_{2n}\left(g_a, X_{n,a}, X_{n,-a}; \theta\right) - 1\left\{G_{n,a} = g_a\right\} \middle| X_{n,a}, X_{n,-a}\right] &\leq 0 \end{aligned} \tag{15}$$

for all g_a, all $(X_{n,a}, X_{n,-a})$, and all $a \leq \bar{a}$. Note that the bounds are invariant under permutations over $[a]^c = [n] \setminus [a]$, so $X_{n,-a}$ can be replaced by its empirical distribution $\phi_n\left(X_{n,-a}\right)$ without information loss. This substantially reduces the dimension of the conditioning variables and prevents it from increasing in n.

We further transform the conditional moment inequalities into equivalent unconditional moment inequalities of the form

$$\begin{aligned} \mathbb{E}\left[\left(1\left\{G_{n,a} = g_a\right\} - H_{1n}\left(g_a, X_{n,a}, \phi_n\left(X_{n,-a}\right); \theta\right)\right) q\left(X_{n,a}, \phi_n\left(X_{n,-a}\right)\right)\right] &\leq 0 \\ \mathbb{E}\left[\left(H_{2n}\left(g_a, X_{n,a}, \phi_n\left(X_{n,-a}\right); \theta\right) - 1\left\{G_{n,a} = g_a\right\}\right) q\left(X_{n,a}, \phi_n\left(X_{n,-a}\right)\right)\right] &\leq 0 \end{aligned} \tag{16}$$

for all nonnegative functions $q\left(X_{n,a}, \phi_n\left(X_{n,-a}\right)\right) \in \mathcal{Q}$, where q represents instruments that depend on the conditioning variables and $\mathcal{Q}$ is a collection of instruments. For discrete X, we can choose $\mathcal{Q} = \left\{1\left\{X_{n,a} = x_a\right\} \cdot \frac{1}{n-a} \sum_{j \notin [a]} 1\left\{X_{n,j} = x\right\} : \forall x_a \in \mathcal{X}_a, \forall x \in \mathcal{X}\right\}$ (with abuse of notation denote $X_{n,a} = \left(X_{n,i}\right)_{i \in [a]}$). If X is continuous, we follow Andrews and Shi (2013) and choose $\mathcal{Q}$ to be a countable set whose elements approximate nonnegative q well, so there is no information loss in the unconditional moment inequalities. For example, we can transform each $X_{n,i}$ to lie in $[0,1]^{d_x}$ and choose $\mathcal{Q}$ to be a collection of indicator functions of cubes in $[0,1]^{d_x}$ with side lengths decreasing to 0, e.g., $\mathcal{Q} = \left\{1\left\{X_{n,a} \in C_a\right\} \cdot \frac{1}{n-a} \sum_{j \notin [a]} 1\left\{X_{n,j} \in C\right\} : \forall C_a \in \mathcal{C}_a, \forall C \in \mathcal{C}\right\}$, where $\mathcal{C} = \left\{\bigotimes_{d=1}^{d_x} \left(\frac{k_d - 1}{2r}, \frac{k_d}{2r}\right] : 1 \leq k_d \leq 2r, 1 \leq d \leq d_x, r = r_0, r_0 + 1, \ldots\right\}$ for some positive

integer r_0, and $\mathcal{C}_a = \bigotimes_{i=1}^{a} \mathcal{C}$. In practice, if $\mathcal{Q}$ is infinite, we approximate it by a finite set via truncation or simulation. See Andrews and Shi (2013) for more details. Note that given the choice of the instruments the unconditional moment inequalities contain terms of the form $1\{G_{n,a} = g_a\} 1\{X_{n,a} = x_a\}$ or $1\{G_{n,a} = g_a\} 1\{X_{n,a} \in C_a\}$. These indicator functions are evaluated in the sense of isomorphism. That is, for a given (g_a, x_a) or (g_a, C_a), the "=" and/or "$\in$" relations hold if they hold for any network obtained from $(G_{n,a}, X_{n,a})$ by a permutation over $[a]$.

The sample moments can be constructed using all subnetworks of size a. In particular, let $A_1, A_2, \ldots, A_{\binom{n}{a}}$ be equal to the $\binom{n}{a}$ subsets of size a of $[n]$. For $t = 1, \ldots, T$, we define the sample moments for the network (G_{n_t}, X_{n_t}) as

$$\begin{aligned} m_1(\theta; G_{n_t}, X_{n_t}, g_a, q) &= \frac{1}{\binom{n_t}{a}} \sum_{i=1}^{\binom{n_t}{a}} \left[\left(1\{G_{n_t,A_i} = g_a\} - H_{1n_t}\left(g_a, X_{n_t,A_i}, \phi_{n_t}(X_{n_t,-A_i}); \theta\right) \right) \cdot \right. \\ &\qquad \left. q\left(X_{n_t,A_i}, \phi_{n_t}(X_{n_t,-A_i})\right) \right] \\ m_2(\theta; G_{n_t}, X_{n_t}, g_a, q) &= \frac{1}{\binom{n_t}{a}} \sum_{i=1}^{\binom{n_t}{a}} \left[\left(H_{2n_t}\left(g_a, X_{n_t,A_i}, \phi_{n_t}(X_{n_t,-A_i}); \theta\right) - 1\{G_{n_t,A_i} = g_a\} \right) \cdot \right. \\ &\qquad \left. q\left(X_{n_t,A_i}, \phi_{n_t}(X_{n_t,-A_i})\right) \right] \end{aligned} \tag{17}$$

for all $g_a \in \mathcal{G}_a$ and all $q \in \mathcal{Q}$. These are valid moments because by exchangeability $\mathbb{E}m_1(\theta; G_{n_t}, X_{n_t}, g_a, q) \leq 0$ and $\mathbb{E}m_2(\theta; G_{n_t}, X_{n_t}, g_a, q) \leq 0$ for $\theta \in \Theta_I(a)$.[19]

The estimation and inference of the identified set are a straightforward application of the moment inequality literature. Details are available upon request.

6 Computation

In this section we discuss how to compute the bounds in (10). Recall that the upper bound is the probability that a subnetwork is PS for some PS complement, and the lower bound has a similar probability form. Computing the events in these probabilities by brute force (e.g., checking all possible complements) is typically infeasible because the number of possible complements is very large even for a moderate n. We propose a more sophisticated method to compute the bounds that is feasible for large

[19] In applications, the number of such subsets N_a may be large if n is large. It suffices to consider a smaller number of randomly chosen subsets of size a, where each subset is obtained by drawing a nodes without replacement uniformly at random from $[n]$.

n. In the sequel we focus on TU. The case of NTU can be handled similarly but with higher computational costs.[20]

Our idea comes from the fact that the bounds can be equivalently represented as functions of certain maximum and minimum marginal utilities over all PS complements. Because the pairwise stability of a complement can be represented by a set of constraints, the maximum and minimum marginal utilities can be computed by solving constrained optimization problems.

To describe the method precisely, let us introduce some notation. For simplicity we suppress the subscript n throughout the section. For any $i < j$, denote by $\Delta V_{ij}(g_{-ij}, x_{ij})$ the sum of i and j's marginal utilities from link ij that depend on the complement g_{-ij} and attributes x_{ij},

$$\Delta V_{ij}(g_{-ij}, x_{ij}) = u(x_{ij}) + u(x_{ji}) + \frac{1}{n-2}\sum_{k\neq i,j}(g_{ik}+g_{jk})\gamma_1 + \frac{2}{n-2}\sum_{k\neq i,j} g_{ik}g_{jk}\gamma_2. \tag{18}$$

Let $\bar{\varepsilon}_{ij} = \varepsilon_{ij} + \varepsilon_{ji}$ for all $i < j$ and $\bar{\varepsilon} = (\bar{\varepsilon}_{ij})_{i<j}$. With abuse of notation let $\mathcal{PS}(x, \bar{\varepsilon})$ denote the collection of PS networks and $\mathcal{PS}(g_{12}, x, \bar{\varepsilon}_{-12})$ the collection of PS complements for link g_{12}, where $\bar{\varepsilon}_{-12} = (\bar{\varepsilon}_{ij})_{(i,j)\neq(1,2)}$ is the preference profile in the complement. Moreover, let $g_{a,-12} = (g_{ij})_{i<j\leq a,\ (i,j)\neq(1,2)}$. Note that $g = (g_{12}, g_{a,-12}, g_{-a})$.

We focus on the upper bound. The lower bound can be handled similarly and details are presented in the appendix. Recall that the upper bound is given by

$$H_{1n}(g_a, x) = \int 1\{\exists g_{-a}, (g_a, g_{-a}) \in \mathcal{PS}(x, \bar{\varepsilon})\} dF(\bar{\varepsilon})$$

It can be equivalently represented as

$$H_{1n}(1, g_{a,-12}, x) = \int 1\{\max_{\substack{g_{-a},\ s.t.\\ (g_{a,-12}, g_{-a})\in\mathcal{PS}(1,x,\bar{\varepsilon}_{-12})}} \Delta V_{12}(g_{a,-12}, g_{-a}, x_{12}) + \bar{\varepsilon}_{12} \geq 0\} dF(\bar{\varepsilon}_{12}, \bar{\varepsilon}_{-12}) \tag{19}$$

for all $g_a = (1, g_{a,-12})$, and

$$H_{1n}(0, g_{a,-12}, x) = \int 1\{\min_{\substack{g_{-a},\ s.t.\\ (g_{a,-12}, g_{-a})\in\mathcal{PS}(0,x,\bar{\varepsilon}_{-12})}} \Delta V_{12}(g_{a,-12}, g_{-a}, x_{12}) + \bar{\varepsilon}_{12} < 0\} dF(\bar{\varepsilon}_{12}, \bar{\varepsilon}_{-12}) \tag{20}$$

[20] The computational cost in NTU for n individuals is approximately that in TU for $\sqrt{2}n$ individuals because i and j's proposals for link ij need to be computed separately.

for all $g_a = (0, g_{a,-12})$, where the max and min are taken over g_{-a}. To understand these expressions, observe that the subnetwork $(1, g_{a,-12})$ is PS for some PS complement g_{-a} if and only if the sum of $\bar{\varepsilon}_{12}$ and the maximum value of $\Delta V_{12}(g_{a,-12}, g_{-a}, x_{12})$ that can be achieved by any PS complement of link g_{12} taking the form $(g_{a,-12}, g_{-a})$ is greater than 0. A similar property holds for the subnetwork $(0, g_{a,-12})$ except that it is the sum of $\bar{\varepsilon}_{12}$ and the minimum value of $\Delta V_{12}(g_{a,-12}, g_{-a}, x_{12})$ that is less than 0. For a given $\bar{\varepsilon}_{-12}$, the maximum or minimum of $\Delta V_{12}(g_{a,-12}, g_{-a}, x_{12})$ gives the boundary of the region where the subnetwork g_a can be part of a PS network, projected in the direction of $\bar{\varepsilon}_{12}$ at the given $\bar{\varepsilon}_{-12}$. Hence, by simulating $\bar{\varepsilon}_{-12}$ we can use these boundary values of $\bar{\varepsilon}_{12}$ to compute the aforementioned region and thus the upper bound.

The optimization problems in (19) and (20) can be further simplified under the utility specification in (1). This specification has a feature that the marginal utility $\Delta V_{ij}(g_{-ij}, x_{ij})$ in (18) depends on the complement g_{-ij} only through the local links that connect i or j to others. We define the collection of such links in the complement g_{-ij} as the *neighborhood* of the pair (i, j), denoted by $b_{ij} = (g_{ik}, g_{jk})_{k \neq i,j}$. More generally, for any set $[a]$ we define the *neighborhood* of the set $[a]$, denote by b_a, as the collection of links in the complement g_{-a} that connect a node in $[a]$ to another outside of $[a]$. Clearly, $g_{-a} = (b_a, g_{a^c})$, where g_{a^c} is the subnetwork in $[a]^c$. For a given subnetwork g_a and its complement g_{-a}, any pair (i, j) in $[a]$ has a neighborhood b_{ij} that depends on the subnetwork g_a and the neighborhood of $[a]$, b_a, but not the subnetwork g_{a^c}, i.e., $b_{ij} = b_{ij}(g_{a,-ij}, b_a)$.[21] Therefore, the maximization problem in (19) can be expressed as

$$\max_{b_a, g_{a^c}} \Delta V_{12}(b_{12}(g_{a,-12}, b_a), x_{12}) \tag{21}$$

$$s.t.\ g_{ij} = 1\{\Delta V_{ij}(b_{ij}(1, g_{a,-12}, b_a), x_{ij}) + \bar{\varepsilon}_{ij} \geq 0\},\ i < j \leq a, (i,j) \neq (1,2) \tag{22}$$

$$g_{ik} = 1\{\Delta V_{ik}(b_{ik}(1, g_{a,-12}, b_a, g_{a^c}), x_{ik}) + \bar{\varepsilon}_{ik} \geq 0\},\ i \leq a, k > a \tag{23}$$

$$g_{kl} = 1\{\Delta V_{kl}(b_{kl}(b_a, g_{a^c}), x_{kl}) + \bar{\varepsilon}_{kl} \geq 0\},\ a < k < l \tag{24}$$

and the minimization problem in (20) has a similar expression, with the max replaced by min and $(1, g_{a,-12})$ replaced by $(0, g_{a,-12})$. The restriction that the complement $(g_{a,-12}, b_a, g_{a^c})$ is PS is represented by the equalities in (22)-(24). This optimization

[21] Similarly, any pair (k, l) not in $[a]$ has a neighborhood b_{kl} that depends on the subnetwork g_{a^c} and the neighborhood of $[a]$, b_a, but not the subnetwork g_a.

problem can be transformed into a linear integer programming problem over b_a and g_{a^c}, from which we can solve for the optimal value of $\Delta V_{12}(b_{12}, x_{12})$.[22]

In practice, because the subnetwork g_{a^c} does not enter the objective function in (21), we can reduce the dimension of the optimization problem by implementing it iteratively over b_a and g_{a^c}. In a special case where links are strategic complements, i.e., $\gamma_1, \gamma_2 \geq 0$, the maximization in (21)-(24) can be solved equivalently by replacing the subnetwork g_{a^c} with the largest PS subnetwork in $[a]^c$ for each given b_a and maximizing the objective function over b_a.[23] We can solve this equivalent problem by iterating over b_a.[24] The same procedure applies to the minimization problem except that the subnetwork g_{a^c} is replaced by the smallest PS subnetwork in $[a]^c$. By iteration we reduce the optimization problems to be over b_a only, which are easy to solve given that the dimension of b_a is $a(n-a)$.[25] Since the spillover effect of b_a on the links in g_{a^c} is small, this iterative procedure tends to converge fast.

In the appendix, we provide more details about the computational procedure and how to handle the general case without strategic complementarity. We also propose a GHK simulator that is smoother than the crude frequency simulators to reduce the simulation error (Hajivassiliou and Ruud (1994), Geweke and Keane (2001)). The idea is to simulate $\bar{\varepsilon}_{-12}$ and solve the optimization problem in (21)-(24) sequentially for each link in $[a]$. Details of the GHK simulator can be found in the appendix.

[22] The optimization problem in (21)-(24) is not fully linear because of (i) the interaction terms of the form $g_{ik}g_{jk}$ in the marginal utilities and (ii) the indicator restrictions in (22)-(24). Nevertheless, we can apply the linearization techniques in integer programming to fully linearize this problem. In particular, for (i) we can introduce an additional binary variable $y = g_{ik}g_{jk}$ for each $g_{ik}g_{jk}$ with the additional inequalities $y \leq g_{ik}$, $y \leq g_{jk}$, and $y \geq g_{ik} + g_{jk} - 1$. As for (ii), an indicator restriction $g_{ij} = 1\{\Delta V_{ij} + \varepsilon_{ij} \geq 0\}$ is equivalent to the linear inequalities $L(1 - g_{ij}) \leq \Delta V_{ij} + \varepsilon_{ij} < M g_{ij}$ for sufficiently large M and sufficiently small L.

[23] When links are strategic complements, the network formation game is a supermodular game (Milgrom and Robert, 1990), and any collection of PS networks has the largest and smallest elements, i.e., there exist PS networks g^0 and g^1 such that $g^0 \leq g \leq g^1$ for any PS network g. The largest and smallest PS networks can be computed from the best-response dynamics (Topkis, 1979).

[24] The iterative procedure consists of the following steps: (i) choose an initial b_a, (ii) compute the largest PS g_{a^c} for $g_{a,-12}$ and the initial b_a, (iii) solve for the optimal b_a that maximizes the objective function under the largest PS g_{a^c}, (iv) update the initial b_a with the optimal b_a, and (v) iterate.

[25] In our simulations, solving such a linear integer programming problem by CPLEX for $n = 100$ and $a = 3$ takes only 0.007 seconds (on a 3.4GHz CPU).

7 Monte Carlo Simulations

In this section, we conduct Monte Carlo simulations to evaluate the subnetwork approach developed in the previous sections. We are in particular interested in the performance of the subnetwork bounds in large networks. Throughout the simulations we consider the marginal utility specification

$$\Delta U_{ij}\left(G_{-ij}, X_i, X_j, \varepsilon_{ij}\right) = \beta\left|X_i - X_j\right| + \frac{1}{n-2}\sum_{k\neq i,j} G_{ik}G_{jk}\gamma + \varepsilon_{ij}$$

where X_i, $i = 1, \ldots, n$, are i.i.d. binary variables that equal to 1 or 0 with equal probability, and ε_{ij}, $i, j = 1, \ldots, n$, $i \neq j$, are i.i.d. $N(0,1)$. The parameter of interest is $\theta = (\beta, \gamma)$. We set the true $\beta_0 = -1$ to create homophily, and assume $\gamma \geq 0$, so the links are strategic complements. We consider pairwise stability in TU. To generate a sample of networks, we compute the largest and smallest PS networks in each observation using the best-response dynamics[26] and let half of the networks in the sample be the largest PS networks and another half the smallest PS networks.

We first investigate the properties of the subnetwork bounds as the network size n increases. To this end, we fix the size of subnetworks at $a = 2$ and consider a variety of network sizes $n = 10, 25, 50, 100, 250, 500$. For each n and each $\gamma \in [0, 3]$, we compute the upper and lower bounds for the subnetwork choice probabilities $\Pr\left(G_{12} = 1 \middle| X_1 = 1, X_2 = 1\right)$ and $\Pr\left(G_{12} = 1 \middle| X_1 = 0, X_2 = 1\right)$ with 100 simulations. The bounds are plotted in Figure 3.

We can see in Figure 3 that the upper and lower bounds are nontrivial for all network sizes. The bounds are close to 1 only for large γ (e.g., $\gamma \geq 2.5$) when the utility externality from friends in common is huge. For such γ, we expect the networks to be complete, so it is reasonable to get close-to-one bounds. The lowest bounds are achieved at $\gamma = 0$ when there is no externality. In this case, the networks coincide with the Erdős-Rényi random graphs with link probability 0.5 for pairs with $X_i = X_j$ and $\Phi\left(-\sqrt{2}\right) = 0.079$ for pairs with $X_i \neq X_j$. The bounds we compute are consistent with these link probabilities. Moreover, all the upper and lower bounds tend to converge as $n \to \infty$. The changes in the bounds become negligible when $n \geq 100$ for all γ. These

[26] Under strategic complementarity, the best-response dynamics converge to the largest PS network if the initial network is chosen to be the largest possible network (e.g. the complete network) and converge to the smallest PS network if the initial network is the smallest possible network (e.g. the empty network).

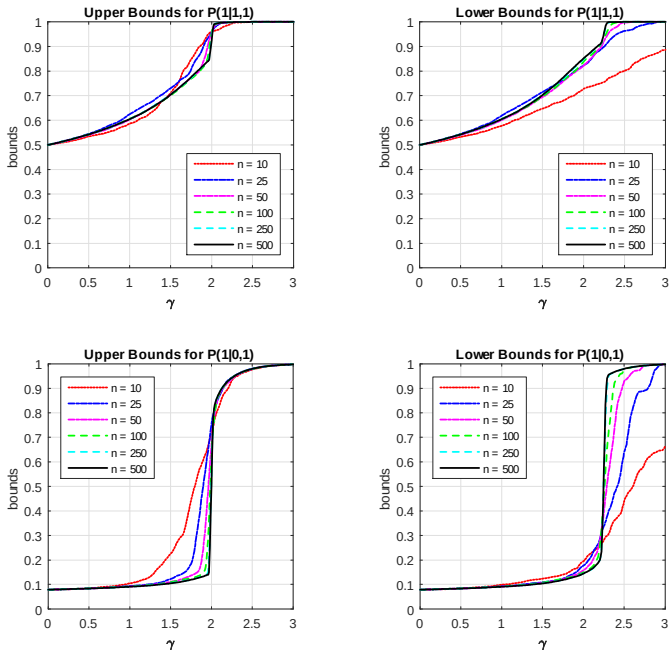

Figure 3: Bounds for Subnetwork Choice Probabilities

results suggest that the bounds provide nontrivial restrictions in large networks.

In addition, note that the upper and lower bounds in Figure 3 are not flat in γ for all network sizes. Since link probabilities lie between these bounds, it is clear that the link probabilities for $\gamma > 0$ are different from those for $\gamma = 0$. This implies that networks generated under our utility specification differ fundamentally from those without network effects, for all network sizes. The network effects do not vanish as networks grow large.

Next we examine whether the subnetwork bounds are informative about the parameter. We set the true $\gamma_0 = 1$ and generate i.i.d. networks of sizes $n = 25, 50, 100$ with sample sizes $T = 50, 200$. For each sample, we consider the bounds from subnetworks of sizes $a = 2$ and $a = 3$ and estimate the corresponding identified sets. We compute the bounds using the methods described in Section 6 with 50 simulations, and construct the sample moments in (17) using 1000 randomly selected subnetworks. For $a = 3$, we also use a graph isomorphism algorithm to determine whether subnetworks are isomorphic.[27] The identified sets are computed using the simulation

[27]We use the graph isomorphism algorithm named Nauty developed by Brendan McKay (http://cs.anu.edu.au/~bdm/nauty). It can calculate isomorphisms for vertex-colored graphs. We

Table 1: Projections of the Estimated Identified Sets

		$a=2$		$a=3$	
T	n	β	γ	β	γ
50	25	$[-1.107, -0.921]$	$[0.813, 1.137]$	$[-1.071, -0.934]$	$[0.868, 1.136]$
		$([-1.205, -0.868])$	$([0.538, 1.338])$	$([-1.115, -0.903])$	$([0.728, 1.251])$
	50	$[-1.107, -0.915]$	$[0.787, 1.129]$	$[-1.069, -0.937]$	$[0.864, 1.123]$
		$([-1.190, -0.860])$	$([0.575, 1.264])$	$([-1.105, -0.912])$	$([0.770, 1.208])$
	100	$[-1.101, -0.917]$	$[0.806, 1.138]$	$[-1.072, -0.937]$	$[0.863, 1.123]$
		$([-1.163, -0.876])$	$([0.621, 1.259])$	$([-1.106, -0.910])$	$([0.772, 1.192])$
200	25	$[-1.104, -0.919]$	$[0.807, 1.126]$	$[-1.071, -0.934]$	$[0.868, 1.132]$
		$([-1.181, -0.866])$	$([0.576, 1.308])$	$([-1.111, -0.908])$	$([0.765, 1.229])$
	50	$[-1.106, -0.915]$	$[0.794, 1.133]$	$[-1.070, -0.937]$	$[0.866, 1.126]$
		$([-1.190, -0.867])$	$([0.565, 1.262])$	$([-1.101, -0.911])$	$([0.771, 1.194])$
	100	$[-1.100, -0.917]$	$[0.808, 1.137]$	$[-1.072, -0.936]$	$[0.859, 1.126]$
		$([-1.162, -0.873])$	$([0.616, 1.261])$	$([-1.107, -0.910])$	$([0.766, 1.198])$
DGP		-1	1	-1	1

Notes: Intervals not in parentheses are the averages of the projections of the identified sets. Intervals in parentheses are the 5% and 95% percentiles of the projections. T is the sample size, n is the network size and a is the subnetwork size.

method suggested by Kline and Tamer (2015).[28] Each identified set is approximated by 100 draws. All the experiments are repeated independently 200 times.

The estimated identified sets are two-dimensional. For each of them, we calculate its one-dimensional projections, i.e., the maximum and minimum values of the simulated β and γ. Then we pool these maxima and minima from the 200 repetitions of each experiment, and calculate their averages, 5% percentiles of the minima, and 95% percentiles of the maxima. These numbers are reported in Table 1 as the mean estimates and confidence intervals for the one-dimensional projections of the identified sets. Moreover, for each experiment we also calculate the values of θ that are

transform a subnetwork (g_a, x_a) into a vertex-colored graph, where the colors of the vertices are defined by x_a, so Nauty is applicable.

[28] In particular, for an identified set defined as $\Theta_I = \{\theta \in \Theta : Q(\theta) = 0\}$ for some function $Q \geq 0$, we simulate random variables from a density proportional to $f_{\Theta_I,\rho}(\theta) = \exp\left(-\frac{Q(\theta)}{\rho}\right)$, where $\rho > 0$ is a small tuning parameter (we choose $\rho = 10^{-4}$) and use the support of the simulated values to approximate the identified set. We implement the simulations by slice sampling (Neal, 2003).

covered by the unions of 90%, 95%, or 99% of the 200 estimated identified sets, and plot them in Figure 4 as the 90%, 95%, and 99% confidence regions of the identified set. Figure 4 is for $T = 50$. The graphs for $T = 200$ are almost identical and are omitted.

From Table 1 and Figure 4 we can see that the bounds from small subnetworks provide informative estimates for the parameter in all the experiments. In particular, the estimates remain stable when we increase the size of the networks. These results are consistent with our earlier findings in Figure 3 and suggest that the subnetwork bounds are informative about the parameter regardless of the network sizes.

Moreover, Table 1 shows that bounds from triples ($a = 3$) are more informative than those from pairs ($a = 2$). For example, the upper bounds of β and the lower bounds of γ become tighter in all the mean estimates and confidence intervals when we move from pairs to triples. The same pattern is observed in the confidence regions in Figure 4. These findings suggest that larger subnetworks can provide more information about the parameter, though the improvement seems to be small.

In addition, we find in Table 1 that the estimates in small samples ($T = 50$) are almost identical to those in large samples ($T = 200$). Averaging over a large number of subnetworks seems to improve the finite sample performance.

8 Conclusion

In this paper, we develop a structural model of network formation. We characterize network formation as a simultaneous-move game, where the decision of forming a link may depend on the linking decisions of others due to utility externalities from indirect friends. With the prevalence of multiple equilibria, the parameters are not necessarily point identified. We propose a partial identification approach that is computationally feasible in large networks. We derive bounds on the probability of observing a subnetwork. These subnetwork bounds are computationally tractable in large networks provided we consider small subnetworks. We provide both theoretical and Monte Carlo evidence that the bounds from small subnetworks are informative about the parameters in large networks.

This subnetwork approach provides a useful framework for exploring the formation of large networks. By focusing on limited aspects of a network rather than solving the full network at once, we can reduce the dimensionality of the problem and ease

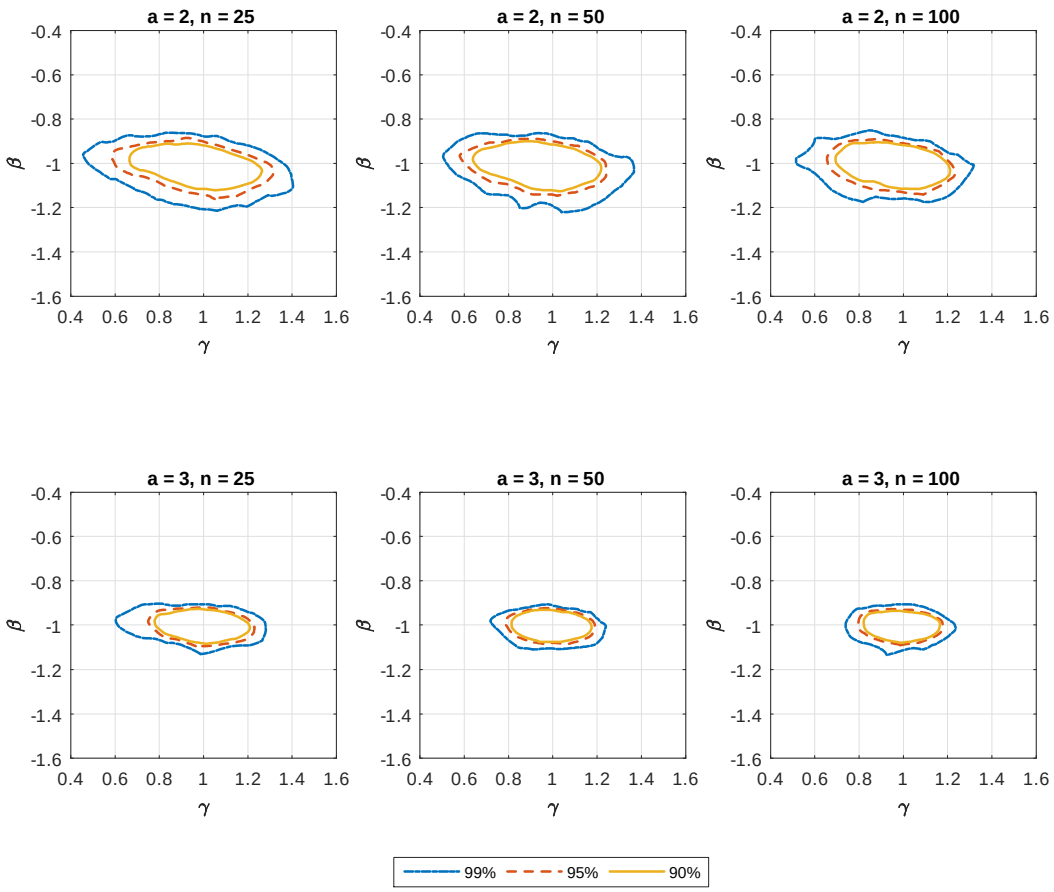

Figure 4: Confidence Regions for the Identified Sets

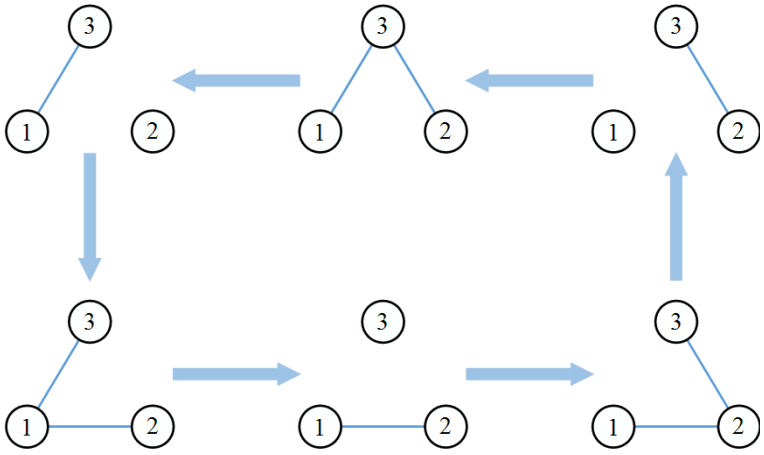

Figure 5: An Example of a Closed Cycle

the computational burden. The spillover effects from the rest of the network are partly internalized in the subnetwork inequalities, so bounds from small subnetworks are able to carry information about the parameters. It may be possible to extend our approach to more general settings that are not covered in the present paper. For example, the networks we consider in the paper are dense. It may be of interest to see whether and how our approach can be extended to networks that are sparse. Another interesting extension is to relate our approach to the literature on large networks and investigate under what conditions the inference based on subnetworks from a single large network is asymptotically valid. These extensions are left for future research.

9 Appendix

9.1 Non-existence of Pairwise Stable Networks

Here we give an example where there is no PS networks, but a closed cycle.

Example 9.1 *Consider networks of size $n = 3$. Suppose the utility function is as in (1) with $u(X_i, X_j; \beta) = 0$, $\gamma_1 < 0$, $\gamma_2 > 0$, $\gamma_1 + \gamma_2 < 0$. Consider the case of NTU. For $\varepsilon_{21}, \varepsilon_{32}, \varepsilon_{13} \geq -\gamma_1$ and $0 \leq \varepsilon_{12}, \varepsilon_{23}, \varepsilon_{31} < -\gamma_1 - \gamma_2$, there is no PS network, but a closed cycle (see Figure 5).*

9.2 Proofs in Section 2

Proof of Proposition 2.1. By Theorem 1 in Jackson and Watts (2001), if there is a function $\Pi : \mathcal{G} \to \mathbb{R}$ such that for any G, G' that differ by one link, G' defeats G if and only if $\Pi(G') > \Pi(G)$, then there is no cycle and thus no closed cycle.[29] In the case of TU, G' defeating G means that for any $i \neq j$ such that $G'_{ij} \neq G_{ij}$, $U_i(G') + U_j(G') > U_i(G) + U_j(G)$. Hence, the proof is complete if we can find such a Π for the utility function in (1).

We show that

$$\Pi(G) = \sum_{i=1}^{n}\sum_{j=1}^{n} G_{ij}u_{ij} + \frac{1}{2\,(n-2)}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\substack{k=1\\k\neq i}}^{n} G_{ij}G_{jk}\gamma_1 + \frac{1}{3\,(n-2)}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} G_{ij}G_{ik}G_{jk}\gamma_2$$

has the desired property, where $u_{ij} = u\,(X_i, X_j; \beta) + \varepsilon_{ij}$. Consider G and G' which differ by link ij. Assume without loss of generality that $G = (0, G_{-ij})$ and $G' = (1, G_{-ij})$. It suffices to show that $\Pi(G') - \Pi(G) = \Delta U_{ij}\,(G_{-ij}) + \Delta U_{ji}\,(G_{-ji})$. By simple algebra

$$\Pi(G') - \Pi(G) = u_{ij} + u_{ji} + \frac{1}{n-2}\sum_{\substack{k=1\\k\neq i,j}}^{n} G_{jk}\gamma_1 + \frac{1}{n-2}\sum_{\substack{k=1\\k\neq i,j}}^{n} G_{ik}\gamma_1 + \frac{2}{n-2}\sum_{\substack{k=1\\k\neq i,j}}^{n} G_{ik}G_{jk}\gamma_2.$$

Moreover, from (2) we have

$$\begin{aligned}
\Delta U_{ij}\,(G_{-ij}) &= u_{ij} + \frac{1}{n-2}\sum_{\substack{k=1\\k\neq i,j}}^{n} G_{jk}\gamma_1 + \frac{1}{n-2}\sum_{\substack{k=1\\k\neq i,j}}^{n} G_{ik}G_{jk}\gamma_2 \\
\Delta U_{ji}\,(G_{-ji}) &= u_{ji} + \frac{1}{n-2}\sum_{\substack{k=1\\k\neq i,j}}^{n} G_{ik}\gamma_1 + \frac{1}{n-2}\sum_{\substack{k=1\\k\neq i,j}}^{n} G_{jk}G_{ik}\gamma_2.
\end{aligned}$$

Hence $\Pi(G') - \Pi(G) = \Delta U_{ij}\,(G_{-ij}) + \Delta U_{ji}\,(G_{-ji})$. The proof is complete. ■

Proof of Proposition 2.2. According to Theorem 1 in Hellmann (2012), if a utility function satisfies convexity in one's own links and strategic complementarity, then there is no closed cycle. A utility function U_i satisfies *convexity in one's own links* if for any $j \neq i$ and $G_{-ij}, G'_{-ij} \in \mathcal{G}_{-ij}$ such that $G_{-ij} = G'_{-ij}$ except that

[29] A *cycle* is a collection of two or more distinct networks that satisfy condition (i) in the definition of closed cycles.

$(G_{-ij})_{ik} = 0$ and $\left(G'_{-ij}\right)_{ik} = 1$ for some $k \neq j$, we have $\Delta U_{ij}(G'_{-ij}) \geq \Delta U_{ij}(G_{-ij})$. In other words, if G'_{-ij} differ from G_{-ij} by adding some links that involve i, the marginal utility of i from link ij with these additional links is larger than without. Moreover, U_i satisfies *strategic complementarity* if for any $j \neq i$ and $G_{-ij}, G'_{-ij} \in \mathcal{G}_{-ij}$ such that $G_{-ij} = G'_{-ij}$ except that $(G_{-ij})_{kl} = 0$ and $\left(G'_{-ij}\right)_{kl} = 1$, for some $k, l \neq i$, we have $\Delta U_{ij}(G'_{-ij}) \geq \Delta U_{ij}(G_{-ij})$. In other words, if G'_{-ij} differ from G_{-ij} by adding some links that do not involve i, the marginal utility of i from link ij given these additional links is larger than without. It suffices to verify that the stated utility function satisfies both properties.

The marginal utility (2) is

$$\Delta U_{ij}\left(G_{-ij}\right) = u_{ij} + \frac{1}{n-2}\sum_{k\neq i,j} G_{jk}\gamma_1 + \frac{1}{n-2}\sum_{k\neq i,j} G_{ik}G_{jk}\gamma_2.$$

where $u_{ij} = u\left(X_i, X_j; \beta\right) + \varepsilon_{ij}$. Since $\gamma_1 \geq 0$ and $\gamma_2 \geq 0$, changing G_{ik} or G_{jk} from 0 to 1 for some $k \neq i$, j weakly increases $\Delta U_{ij}\left(G_{-ij}\right)$. Hence both properties are satisfied. The proof is complete. ■

9.3 Proofs in Section 4

Proof of Proposition 4.2. Without loss of generality we assume TU, and the case of NTU can be proved similarly. For any $i < j \leq n$, define $v_{ij}\left(g_{n,-ij}\right)$ to be the marginal utility of i forming a link with j that is due to the utility externality from other links, i.e.,

$$v_{ij}\left(g_{n,-ij}\right) = \frac{1}{n-2}\sum_{k\neq i,j} g_{n,jk}\gamma_1 + \frac{1}{n-2}\sum_{k\neq i,j} g_{n,ik}g_{n,jk}\gamma_2.$$

Since both $\frac{1}{n-2}\sum_{k\neq i,j} g_{n,jk}$ and $\frac{1}{n-2}\sum_{k\neq i,j} g_{n,ik}g_{n,jk}$ are bounded between 0 and 1, there exist finite constants v^l and v^l such that $v^l \leq v_{ij}\left(g_{n,-ij}\right) \leq v^u$ for all $g_{n,-ij}$. Let $\bar{u}\left(x_{ij}\right) = u\left(x_i, x_j\right) + u\left(x_j, x_i\right)$ and $\bar{\varepsilon}_{ij} = \varepsilon_{ij} + \varepsilon_{ji}$.

The upper bound $H_{1n}\left(g_a, x_a, X_{n,-a}\right)$ is the probability that there is a complement $g_{n,-a}$ such that $(g_a, g_{n,-a})$ is pairwise stable. By the definition of pairwise stability for such a $g_{n,-a}$ the sum of the marginal utilities of i and j from link ij for any $i < j \leq a$

satisfies

$$\Delta U_{ij} + \Delta U_{ji} = \bar{u}\left(x_{ij}\right) + v_{ij}\left(g_{n,-ij}\right) + v_{ji}\left(g_{n,-ji}\right) + \bar{\varepsilon}_{ij} \begin{cases} \geq 0, & \text{if } g_{ij} = 1 \\ < 0, & \text{if } g_{ij} = 0 \end{cases} \tag{25}$$

Since the overall network effect is bounded by $2v^l \leq v_{ij}\left(g_{n,-ij}\right) + v_{ji}\left(g_{n,-ji}\right) \leq 2v^u$, the event in (25) implies that $\bar{u}\left(x_{ij}\right) + 2v^u + \bar{\varepsilon}_{ij} \geq 0$ if $g_{ij} = 1$ and $\bar{u}(x_{ij}) + 2v^l + \bar{\varepsilon}_{ij} < 0$ if $g_{ij} = 0$. Hence,

$$\begin{aligned} H_{1n}\left(g_a, x_a, X_{n,-a}\right) \quad \leq \quad & \prod_{\substack{i<j\leq a \\ g_{ij}=1}} \Pr\left(\bar{u}\left(x_{ij}\right) + 2v^u + \bar{\varepsilon}_{ij} \geq 0\right) \\ & \cdot \prod_{\substack{i<j\leq a \\ g_{ij}=0}} \Pr\left(\bar{u}\left(x_{ij}\right) + 2v^l + \bar{\varepsilon}_{ij} < 0\right) := \bar{H}_1\left(g_a, x_a\right). \end{aligned}$$

Let $\bar{H}_1\left(g_a, x_a\right)$ define the right-hand side. It is strictly smaller than 1 because v^u and v^l are bounded.

Similarly, the lower bound $H_{2n}\left(g_a, x_a, X_{n,-a}\right)$ is the probability that there is a complement $g_{n,-a}$ such that $(g_a, g_{n,-a})$ is pairwise stable and only subnetwork g_a has this property. For such a $g_{n,-a}$ the sum of the marginal utilities of i and j from link ij for $i < j \leq a$ also satisfies the event in (25), which holds if $\bar{u}\left(x_{ij}\right) + 2v^l + \bar{\varepsilon}_{ij} \geq 0$ if $g_{ij} = 1$ and $\bar{u}\left(x_{ij}\right) + 2v^u + \bar{\varepsilon}_{ij} < 0$ if $g_{ij} = 0$. Moreover, when this event occurs there is no $g'_a \neq g_a$ that can satisfy the pairwise stability condition. Therefore,

$$\begin{aligned} H_{2n}\left(g_a, x_a, X_{n,-a}\right) \quad \geq \quad & \prod_{\substack{i<j\leq a \\ g_{ij}=1}} \Pr\left(\bar{u}\left(x_{ij}\right) + 2v^l + \bar{\varepsilon}_{ij} \geq 0\right) \\ & \cdot \prod_{\substack{i<j\leq a \\ g_{ij}=0}} \Pr\left(\bar{u}\left(x_{ij}\right) + 2v^u + \bar{\varepsilon}_{ij} < 0\right) := \bar{H}_2\left(g_a, x_a\right). \end{aligned}$$

Let $\bar{H}_2\left(g_a, x_a\right)$ define the right-hand side. It is strictly greater than 0 because of the boundedness of v^u and v^l. ■

9.4 Auxiliary Results in Section 6

In this section, we provide more details about the computation of the bounds. Recall that the upper bound can be represented as in (19) and (20). In practice, we can

construct a simulator for the upper bound by (i) simulating i.i.d. $\bar{\varepsilon}_{-12}$ and (ii) solving for the maximum or minimum value of $\Delta V_{12}\left(g_{a,-12}, b_a, x_{12}\right)$ from the optimization problem in (21)-(24) at each simulated $\bar{\varepsilon}_{-12}$.

To be more specific about the simulator, for a given $\bar{\varepsilon}_{-12}$, denote the maximum and minimum values of $\Delta V_{12}\left(g_{a,-12}, b_a, x_{12}\right)$ in (19) and (20) by $\Delta V_{12}^{\max}\left(g_{a,-12}, x, \bar{\varepsilon}_{-12}\right)$ and $\Delta V_{12}^{\min}(g_{a,-12}, x, \bar{\varepsilon}_{-12})$. Let $F_{\bar{\varepsilon}}$ denote the CDF of $\bar{\varepsilon}_{ij}$. Observe that the upper bound in (19) and (20) can be expressed as

$$H_{1n}\left(1, g_{a,-12}, x\right) = \int \left(1 - F_{\bar{\varepsilon}}\left(-\Delta V_{12}^{\max}\left(g_{a,-12}, x, \bar{\varepsilon}_{-12}\right)\right)\right) dF\left(\bar{\varepsilon}_{-12}\right)$$

for all $g_a = (1, g_{a,-12})$ and

$$H_{1n}\left(0, g_{a,-12}, x\right) = \int F_{\bar{\varepsilon}}\left(-\Delta V_{12}^{\min}(g_{a,-12}, x, \bar{\varepsilon}_{-12})\right) dF\left(\bar{\varepsilon}_{-12}\right)$$

for all $g_a = (0, g_{a,-12})$. Therefore, a simulator for the upper bound can be constructed by taking the average over the values of $1 - F_{\bar{\varepsilon}}(-\Delta V_{12}^{\max}(g_{a,-12}, x, \bar{\varepsilon}_{-12}))$ (or $F_{\bar{\varepsilon}}(-\Delta V_{12}^{\min}(g_{a,-12}, x, \bar{\varepsilon}_{-12}))$) computed at each simulated $\bar{\varepsilon}_{-12}$.

When we solve the optimization problem in (21)-(24) at a simulated $\bar{\varepsilon}_{-12}$, it is possible that the links in $g_{a,-12}$ cannot be PS for any PS complement g_{-a} (i.e., the constraints in (22) can never be satisfied), so the optimization problem has no solution. For such a $\bar{\varepsilon}_{-12}$, since the indicator function in (19) (or (20)) is equal to 0, we can simply set the value of $1 - F_{\bar{\varepsilon}}\left(-\Delta V_{12}^{\max}\left(g_{a,-12}, x, \bar{\varepsilon}_{-12}\right)\right)$ (or $F_{\bar{\varepsilon}}\left(-\Delta V_{12}^{\min}(g_{a,-12}, x, \bar{\varepsilon}_{-12})\right)$) to be 0. The aforementioned simulator is still valid, but it is nonsmooth similarly to standard crude frequency simulators (McFadden (1989), Pakes and Pollard (1989)), thereby requiring a large number of simulations to reduce the simulation error.

Following the GHK algorithm (Hajivassiliou and Ruud (1994), Geweke and Keane (2001)), we propose a sophisticated simulator that is smoother than the aforementioned simulator, so that the number of simulations can be smaller. The idea is to simulate $\bar{\varepsilon}_{a,-12}$, the components of $\bar{\varepsilon}_{-12}$ in $[a]$, sequentially and solve a sequence of optimization problems as in (21)-(24) for each link in $[a]$.

For expositional simplicity, we describe the algorithm for the simulator in an example of $a = 3$. Note that $\bar{\varepsilon}_{-12} = (\bar{\varepsilon}_{a,-12}, \bar{\varepsilon}_{-a}) = (\bar{\varepsilon}_{13}, \bar{\varepsilon}_{23}, \bar{\varepsilon}_{-a})$.

Algorithm 1 *1. For each simulated* $\bar{\varepsilon}_{-a}$*,*

(1) For $g_{23} = 1$ or 0, (a) solve the problem

$$\begin{array}{rl} \max\limits_{b_a, g_a^c} / \min\limits_{b_a, g_a^c} & \Delta V_{23}\left(g_{a,-23}, b_a, x_{23}\right) \\ s.t. & \textit{equalities (23)-(24)} \end{array}$$

respectively, and (b) generate $\bar{\varepsilon}_{23}$ from the conditional distribution

$$\bar{\varepsilon}_{23} \sim \begin{cases} F_{\bar{\varepsilon}}\left(\bar{\varepsilon}_{ij} \middle| \bar{\varepsilon}_{ij} \geq -\Delta V_{23}^{\max}\left(g_{a,-23}, x, \bar{\varepsilon}_{-a}\right)\right), & \textit{if } g_{23} = 1, \\ F_{\bar{\varepsilon}}\left(\bar{\varepsilon}_{ij} \middle| \bar{\varepsilon}_{ij} < -\Delta V_{23}^{\min}\left(g_{a,-23}, x, \bar{\varepsilon}_{-a}\right)\right), & \textit{if } g_{23} = 0. \end{cases}$$

(2) For $g_{13} = 1$ or 0, (a) solve the problem

$$\begin{array}{rl} \max\limits_{b_a, g_a^c} / \min\limits_{b_a, g_a^c} & \Delta V_{13}\left(g_{a,-13}, b_a, x_{13}\right) \\ s.t. & g_{23} = 1\left\{\Delta V_{23}\left(g_{a,-23}, b_a, x_{23}\right) + \bar{\varepsilon}_{23} \geq 0\right\} \\ & \textit{equalities (23)-(24)} \end{array}$$

respectively, and (b) generate $\bar{\varepsilon}_{13}$ from the conditional distribution

$$\bar{\varepsilon}_{13} \sim \begin{cases} F_{\bar{\varepsilon}}\left(\bar{\varepsilon}_{ij} \middle| \bar{\varepsilon}_{ij} \geq -\Delta V_{13}^{\max}\left(g_{a,-13}, x, \bar{\varepsilon}_{23}, \bar{\varepsilon}_{-a}\right)\right), & \textit{if } g_{13} = 1, \\ F_{\bar{\varepsilon}}\left(\bar{\varepsilon}_{ij} \middle| \bar{\varepsilon}_{ij} < -\Delta V_{13}^{\min}\left(g_{a,-13}, x, \bar{\varepsilon}_{23}, \bar{\varepsilon}_{-a}\right)\right), & \textit{if } g_{13} = 0. \end{cases}$$

(3) For $g_{12} = 1$ or 0, solve the problem

$$\begin{array}{rl} \max\limits_{b_a, g_a^c} / \min\limits_{b_a, g_a^c} & \Delta V_{12}\left(g_{a,-12}, b_a, x_{12}\right) \\ s.t. & g_{23} = 1\left\{\Delta V_{23}\left(g_{a,-23}, b_a, x_{23}\right) + \bar{\varepsilon}_{23} \geq 0\right\} \\ & g_{13} = 1\left\{\Delta V_{13}\left(g_{a,-13}, b_a, x_{13}\right) + \bar{\varepsilon}_{13} \geq 0\right\} \\ & \textit{equalities (23)-(24).} \end{array}$$

respectively.

(4) Calculate

$$P_{ij} = \begin{cases} 1 - F_{\bar{\varepsilon}}\left(-\Delta V_{ij}^{\max}\left(g_{a,-ij}, x, \cdot\right)\right), & \textit{if } g_{ij} = 1, \\ F_{\bar{\varepsilon}}\left(-\Delta V_{ij}^{\min}\left(g_{a,-ij}, x, \cdot\right)\right), & \textit{if } g_{ij} = 0. \end{cases}$$

for $i < j \leq 3$, and the value of their product $\prod_{i<j\leq 3} P_{ij}$. The latter gives one simu-

lation of the upper bound.

2. Repeat Step 1 independently R times, and take the average over the R values of $\prod_{i<j\leq 3} P_{ij}$. This average gives the simulator for the upper bound.

In Algorithm 1, because $\bar{\varepsilon}_{23}$ is drawn from a conditional distribution given that the link g_{23} is PS for some PS complement g_{-a}, and $\bar{\varepsilon}_{13}$ is drawn from a conditional distribution given that the links g_{13} and g_{23} are PS for some PS complement g_{-a}, the optimization problems in Steps 1(2)-(3) for such a $(\bar{\varepsilon}_{13}, \bar{\varepsilon}_{23})$ are guaranteed to have a solution. The simulator is then constructed by taking the average over the values of P_{12}, weighted appropriately (i.e., multiplied by $P_{13}P_{23}$) to account for the difference between the conditional distribution and unconditional distribution of $(\bar{\varepsilon}_{13}, \bar{\varepsilon}_{23})$. This simulator is smoother in the parameter than the aforementioned naive simulator.

As for the implementation of the optimization problem in (21)-(24), we have discussed in Section 6 how to solve it in the special case of strategic complementarity. In the general case without strategic complementarity, we deal with the subnetwork g_{a^c} by making use of the property of potential games. Recall that in this general case to ensure the existence of a PS network we need to assume TU so that under our utility specification the game can be represented as a potential game. From the property of potential games a PS network is a local maximum of the potential function, so computing PS subnetworks g_{a^c} amounts to finding local maxima of the potential function. While finding an exact local maximum is a NP problem, it is possible to find an approximate local maximum in polynomial time. For example, Orlin, Punnen and Schulz (2004) show that an ε-local maximum can be found in time polynomial in the problem size and $\frac{1}{\varepsilon}$. Hence, we can solve the optimization problem in (21)-(24) approximately by replacing the equalities in (24) with an availability constraint on b_a, i.e., a neighborhood b_a is available if it is PS for some approximate PS subnetwork g_{a^c}.[30] We expect that the approximation in the PS subnetwork g_{a^c} has a negligible effect on the optimal value because g_{a^c} plays a role only through the availability of b_a, and the effect of a link on the marginal utility of another is at most at the order of $\frac{1}{n-2}$.

Next we consider the lower bound. It is given by

$$H_{2n}(g_a, x)$$

[30] Such problems can be solved using a constraint integer programming solver like SCIP.

$$
\begin{aligned}
&= \int 1\left\{\exists g_{-a},(g_a,g_{-a})\in\mathcal{PS}\left(x,\bar{\varepsilon}\right)\&\forall g'_a\neq g_a,\forall g_{-a},(g'_a,g_{-a})\notin\mathcal{PS}\left(x,\bar{\varepsilon}\right)\right\}dF\left(\bar{\varepsilon}\right)\\
&= 1-\int 1\left\{\exists g'_a\neq g_a,\exists g_{-a},(g'_a,g_{-a})\in\mathcal{PS}\left(x,\bar{\varepsilon}\right)\right\}dF\left(\bar{\varepsilon}\right). \qquad (26)
\end{aligned}
$$

The last equality follows because $\Pr\left(A\cap B^c\right)=\Pr\left(A\cup B\right)-\Pr\left(B\right)$ for the sets $A=\{\bar{\varepsilon}:\exists g_{-a},(g_a,g_{-a})\in\mathcal{PS}\left(x,\bar{\varepsilon}\right)\}$ and $B=\{\bar{\varepsilon}:\exists g'_a\neq g_a,\exists g_{-a},(g'_a,g_{-a})\in\mathcal{PS}\left(x,\bar{\varepsilon}\right)\}$, and $\Pr\left(A\cup B\right)=1$ since the equilibrium set $\mathcal{PS}\left(x,\bar{\varepsilon}\right)$ is not empty. For $a=2$, the bounds satisfy $H_{2n}\left(g_{12}=1,x\right)=1-H_{1n}\left(g_{12}=0,x\right)$ and $H_{2n}\left(g_{12}=0,x\right)=1-H_{1n}\left(g_{12}=1,x\right)$, so we get the lower bound immediately. For $a>2$, note that the indicator function in (26) says that there is a subnetwork $g'_a\neq g_a$ such that g'_a is PS for some PS complement g_{-a}. Such a subnetwork g'_a may have $g'_{12}=1$ or 0. By considering the two cases separately we can represent this indicator function similarly to those in (19) and (20), i.e.,

$$
\begin{aligned}
&1\left\{\exists g'_a\neq g_a,\exists g_{-a},(g'_a,g_{-a})\in\mathcal{PS}\left(x,\bar{\varepsilon}\right)\right\}\\
= \;& 1\{\max_{\substack{g'_{a,-12},g_{-a},\ s.t.\\ (g'_{a,-12},g_{-a})\in\mathcal{PS}(1,x,\bar{\varepsilon}_{-12}),\ (1,g'_{a,-12})\neq g_a}}\Delta V_{12}\left(g'_{a,-12},g_{-a},x_{12}\right)+\bar{\varepsilon}_{12}\geq 0\}\vee\\
& 1\{\min_{\substack{g'_{a,-12},g_{-a},\ s.t.\\ (g'_{a,-12},g_{-a})\in\mathcal{PS}(0,x,\bar{\varepsilon}_{-12}),\ (0,g'_{a,-12})\neq g_a}}\Delta V_{12}\left(g'_{a,-12},g_{-a},x_{12}\right)+\bar{\varepsilon}_{12}<0\} \qquad (27)
\end{aligned}
$$

where $x\vee y=\max\left(x,y\right)$ and the max and min are over $g'_{a,-12}$ and g_{-a}.

For a given ε_{-12}, denote the maximum and minimum values of $\Delta V_{12}\left(g'_{a,-12},g_{-a},x_{12}\right)$ in (27) by $\Delta V_{12}^{\max}(g_a,x,\bar{\varepsilon}_{-12})$ and $\Delta V_{12}^{\min}(g_a,x,\bar{\varepsilon}_{-12})$. The event in (27) occurs if $\bar{\varepsilon}_{12}\geq-\Delta V_{12}^{\max}(g_a,x,\bar{\varepsilon}_{-12})$ or $\bar{\varepsilon}_{12}<-\Delta V_{12}^{\min}(g_a,x,\bar{\varepsilon}_{-12})$. Hence, a simulator for the lower bound can be constructed using the values of the probability that the event in (27) does not occur calculated at each simulated $\bar{\varepsilon}_{-12}$. Similarly to the upper bound, we can construct a GHK simulator for the lower bound as well.

The maximization problem in (27) can be expressed similarly to that in (21)-(24), i.e.,

$$
\begin{aligned}
\max_{g'_{a,-12},b_a,g_{a^c}}\quad & \Delta V_{12}\left(b_{12}\left(g'_{a,-12},b_a\right),x_{12}\right)\\
s.t.\quad & g'_{ij}=1\left\{\Delta V_{ij}\left(b_{ij}\left(1,g'_{a,-12},b_a\right),x_{ij}\right)+\bar{\varepsilon}_{ij}\geq 0\right\},\ \ i<j\leq a,\ \ (i,j)\neq(1,2)\\
& g_{ik}=1\left\{\Delta V_{ik}\left(b_{ik}\left(1,g'_{a,-12},b_a,g_{a^c}\right),x_{ik}\right)+\bar{\varepsilon}_{ik}\geq 0\right\},\ \ i\leq a,\ k>a\\
& g_{kl}=1\left\{\Delta V_{kl}\left(b_{kl}\left(b_a,g_{a^c}\right),x_{kl}\right)+\bar{\varepsilon}_{kl}\geq 0\right\},\ \ a<k<l
\end{aligned}
$$

$$\left(1, g'_{a,-12}\right) \neq g_a.$$

and the minimization problem in (27) has a similar expression, with the max replaced by min and $\left(1, g'_{a,-12}\right)$ replaced by $\left(0, g'_{a,-12}\right)$. These optimization problems can be solved similarly using the aforementioned methods.

References

[1] Andrews, D. W. K., S., Berry, and P. Jia (2004) Confidence Regions for Parameters in Discrete Games with Multiple Equilibria, with an Application to Discount Chain Store Location, working paper.

[2] Andrews, D. W. K., and P. Jia (2012) Inference for Parameters Defined by Moment Inequalities: A Recommended Moment Selection Procedure, *Econometrica*, 80(6): 2805–2826.

[3] Andrews, D. W. K., and X. Shi (2013) Inference Based on Conditinal Moment Inequalities, *Econometrica*, 81(2): 609-666.

[4] Andrews, D. W. K., and G. Soares (2010) Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection, *Econometrica*, 78(1): 119-157.

[5] Bajari, P., J. Hahn, H. Hong, and G. Ridder (2011) A Note on Semiparametric Estimation of Finite Mixtures of Discrete Choice Models with Application to Game Theoretic Models, *International Economic Review*, 53 (3), 807-824.

[6] Bajari, P., H. Hong., and S. P. Ryan (2010) Identification and Estimation of a Discrete Game of Complete Information, *Econometrica*, 78(5): 1529-1568.

[7] Belleflamme, P., and F. Bloch (2004) Market Sharing Agreements and Collusive Networks. *International Economic Review*, 45(2): 387-411.

[8] Beresteanu, A., I. Molchanov, and F. Molinari (2011) Sharp Identification Regions in Models With Convex Moment Predictions, *Econometrica*, 79(6): 1785–1821.

[9] Berry, S., and E. Tamer (2006) Identification in Models of Oligopoly Entry, in *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, Volume II, edited by R. Blundell, W. K. Newey, and T. Persson, Cambridge University Press.

[10] Bhamidi, S., G. Bresler, and A. Sly (2011) Mixing Time of Exponential Random Graphs, *The Annals of Applied Probability*, 21(6): 2146-2170.

[11] Bierlaire, M., D. Bolduc, and D. McFadden (2008) The Estimation of Generalized Extreme Value Models from Choice-based Samples, *Transportation Research Part B: Methodological,* 42(4): 381-394.

[12] Bloch, F., and M. O. Jackson (2006) Definitions of Equilibrium in Network Formation Games, *International Journal of Game Theory*, 34: 305-318.

[13] Bloch, F., and M. O. Jackson (2007) The Formation of Networks with Transfers Among Players, *Journal of Economic Theory* 133: 83-110.

[14] Bresnahan, T. F., and P. C. Reiss (1991) Empirical Models of Discrete Games, *Journal of Econometrics*, 48: 57-81.

[15] Bollobás, B. and O. Riordan (2009) Metrics for Sparse Graphs, arXiv: 0708.1919v3.

[16] Boucher, V., and I. Mourifié (2017) My Friend Far, Far Away: A Random Field Approach to Exponential Random Graph Models, *Econometrics Journal*, 20(3): S14-S46.

[17] Calvó-Armengol, A., and M. O. Jackson (2004) The Effects of Social Networks on Employment and Inequality, *American Economic Review*, 94(3): 426-454.

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

[19] Caron, F. and E. Fox (2015) Sparse Graphs Using Exchangeable Random Meaures, *arXiv: 1401.1137.*

[20] Chandrasekhar, A. G., and M. O. Jackson (2013) Tractable and Consistent Random Graph Models, working paper.

[21] Chernozhukov, V., and H. Hong (2004) Likelihood Estimation and Inference in a Class of Nonregular Econometric Models, *Econometrica*, 72(5), 1445-1480.

[22] Chernozhukov, V., H. Hong, and E. Tamer (2007) Estimation and Confidence Regions for Parameter Sets in Econometric Models, *Econometrica*, 75(5), 1243-1284.

[23] Christakis, N., J. Fowler, G. W. Imbens, and K. Kalyanaraman (2010) An Empirical Model for Strategic Network Formation, NBER working paper No.16039.

[24] Ciliberto, F., and E. Tamer (2009) Market Structure and Multiple Equilibria in Airline Markets, *Econometrica, 77(6)*, 1791-1828.

[25] Conley, T., and C. Udry (2010) Learning About a New Technology: Pineapple in Ghana, *American Economic Review*, 100(1), 35-69.

[26] Currarini S., M. O. Jackson and P. Pin (2009) An Economic Model of Friendship: Homophily, Minorities, and Segregation, *Econometrica*, 77(4): 1003-1045.

[27] de Paula, A., S. Richards-Shubik, and E. Tamer (2018) Identifying Preferences in Networks With Bounded Degree, *Econometrica*, 86(1): 263-288.

[28] Diaconis, P. and S. Janson (2008) Graph Limits and Exchangeable Random Graphs, arXiv: 0712.2749v1.

[29] Dutta, B., and S. Mutuswami (1997) Stable Networks, *Journal of Economic Theory* 76: 322-344.

[30] Erdős, P., and A. Rényi (1959) On Random Graphs I, *Publicationes Mathematicae Debrecen* 6: 290–297.

[31] Fafchamps, M., and F. Gubert (2007) The Formation of Risk Sharing Networks, *Journal of Development Economics*, 83: 326-350.

[32] Geweke, J., and M. Keane (2001) Computationally Intensive Methods for Integration in Econometrics, in J. J. Heckman and E. Leamer eds., *Handbook of Econometrics*, Volume 5, Chapter 56: 3463-3568.

[33] Goyal, S. (2007) *Connections: An Introduction to the Economics of Networks*, Princeton University Press, Princeton, NJ.

[34] Goyal, S., and S. Joshi (2006) Unequal Connections, *International Journal of Game Theory*, 34: 319-349.

[35] Goyal, S., and F. Vega-Redondo (2007) Structural Holes in Social Networks, *Journal of Economic Theory*, 137: 460-492.

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

[37] Graham, B. (2018) Econometric Models of Network Formation, prepared for the *Handbook of Econometrics* (Vol. 7A).

[38] Hajivassiliou, V. A., and P. A. Ruud (1994) Classical Estimation Methods for LDV Models Using Simulation, in R. F. Engle and D. L. McFadden eds., *Handbook of Econometrics*, Volume 4, Chapter 40: 2384-2443.

[39] Hellmann, T. (2012) On the Existence and Uniqueness of Pairwise Stable Networks, *International Journal of Game Theory*, forthcoming.

[40] Hirano, K., and J. R. Porter (2003) Asymptotic Efficiency in Parametric Structural Models with Parameter-Dependent Support, *Econometrica*, 71(5), 1307–1338.

[41] Jackson, M. O. (2008) *Social and Economic Networks*, Princeton University Press, Princeton, NJ.

[42] Jackson, M. O., T. Barraquer, and X. Tan (2012) Social Capital and Social Quilts: Network Patterns of Favor Exchange, *American Economic Review*, 102 (5): 1857-1897.

[43] Jackson, M. O., and A. van der Nouweland (2005) Strongly stable networks, *Games and Economic Behavior*, 51: 420-444.

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

[45] Jackson, M. O., and A. Watts (2001) The Existence of Pairwise Stable Networks, *Seoul Journal of Economics*, 14, 3: 299-321.

[46] Jackson, M. O., and A. Watts (2002) The Evolution of Social and Economic Networks, *Journal of Economic Theory,* 106: 265-295.

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

[48] Kallenberg, O. (2005) *Probabilistic Symmetries and Invariance Principles*, Springer, New York, NY.

[49] Kline, B. and E. Tamer (2015) Bayesian Inference in a Class of Partially Identified Models, working paper.

[50] Leung, M. (2015) Two-Step Estimation of Network-Formation Models with Incomplete Information, *Journal of Econometrics*, 188: 182-195.

[51] Leung, M. (2018) A Weak Law for Moments of Pairwise-Stable Networks, SSRN working paper.

[52] Lovász, L. and B. Szegedy (2006) Limits of Dense Graph Sequences, *Journal of Combinatorial Theory*, Series B 96: 933-957.

[53] Lovász, L. (2012). *Large Networks and Graph Limit*, American Mathematical Society Colloquium Publications, Volumn 60.

[54] Mayer A., and S. L. Puller (2008) The Old Boy (and Girl) Network: Social Network Formation on University Campuses, *Journal of Public Economics*, 92: 329–347.

[55] McFadden, D. (1989) A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration, *Econometrica*, 57(5): 995-1026.

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

[57] Mele, A. and L. Zhu (2017) Approximate Variational Estimation for a Model of Network Formation. Technical report, John Hopkins University.

[58] Menzel, K. (2015) Large Matching Markets as Two-Sided Demand Systems, *Econometrica*, 83(3): 897-941.

[59] Menzel, K. (2016) Inference for Games with Many Players, *Review of Economic Studies*, 83: 306-337.

[60] Menzel, K. (2017) Strategic Network Formation with Many Agents, working paper.

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

[62] Milgrom, P. and J. Roberts (1990) Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities, *Econometrica*, 58(6): 1255-1277.

[63] Moretti, E. (2011) Social Learning and Peer Effects in Consumption: Evidence from Movie Sales, *Review of Economic Studies*, 78, 356-393.

[64] Monderer, D. and L. S. Shapley (1994) Potential Games, *Games and Economic Behavior*, 14: 124-143.

[65] Myerson, R. (1991) *Game Theory: Analysis of Conflict*. Harvard University Press.

[66] Nakajima, R. (2007) Measuring Peer Effects on Youth Smoking Behavior, *Review of Economic Studies*, 74(3): 897-935.

[67] Neal, R. (2003) Slice Sampling, *Annals of Statistics*, 31(3): 705-767.

[68] Orlin J., A. Punnen, and A. Schulz (2004) Approximate Local Search in Combinatorial Optimization, *SIAM Journal on Computing*, 33(5): 1201-1214.

[69] Pakes, A., and D. Pollard (1989) Simulation and the Asymptotics of Optimization Estimators, *Econometrica,* 57(5): 1027-1057.

[70] Pakes, A., J. Porter, K. Ho, and J. Ishii (2006) Moment Inequalities and Their Application, working paper.

[71] Ridder, G., and S. Sheng (2018) Estimation of Large Network Formation Games, working paper.

[72] Romano, J. P., and A. M. Shaikh (2010) Inference for the Identified Set in Partially Identified Econometric Models, *Econometrica*, 78(1): 169-211.

[73] Snijders, T. (2002) Markov Chain Monte Carlo Estimation of Exponential Random Graph Models, *Journal of Social Structure*, 3(2).

[74] Tamer, E. (2003) Incomplete Simultaneous Discrete Response Model with Multiple Equilibria, *Review of Economic Studies*, 70: 147-190.

[75] Topkis, D. M. (1979) Equilibrium Points in Nonzero-Sum n-Person Submodular Games, *SIAM Journal of Control and Optimization*, 17(6): 773-787.

[76] van der Vaart, A., and J. A. Wellner (1996) *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer.

[77] Young, H. P. (1993) The Evolution of Conventions, *Econometrica*, 61(1): 57-84.