cemmap

centre for microdata methods and practice

Identifying preferences in networks with bounded degree

Áureo De Paula
Seth Richards-Shubik
Elie Tamer

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP35/17

An ESRC Research Centre

Identifying Preferences in Networks with Bounded Degree*

Áureo de Paula[†]
UCL, São Paulo School of Economics,
CeMMAP, IFS and CEPR

Seth Richards-Shubik[‡]
Lehigh University
and NBER

Elie Tamer[§]
Harvard University

First Version: December 2011
This Version: August 2017

Abstract

This paper provides a framework for identifying preferences in a large network where links are pairwise stable. Network formation models present difficulties for identification, especially when links can be interdependent: e.g., when indirect connections matter. We show how one can use the observed proportions of various local network structures to learn about the underlying preference parameters. The key assumption for our approach restricts individuals to have bounded degree in equilibrium, implying a finite number of payoff-relevant local structures. Our main result provides necessary conditions for parameters to belong to the identified set. We then develop a quadratic programming algorithm that can be used to construct this set. With further restrictions on preferences, we show that our conditions are also sufficient for pairwise stability and therefore characterize the identified set precisely. Overall, the use of both the economic model along with pairwise stability allows us to obtain effective dimension reduction.

*We thank seminar and conference participants at the Conference on Econometric Analysis of Social Interactions at Northwestern University (2010), the ASSA meetings in Chicago (2012), University of Western Ontario (2013), the Stats in Paris: Statistics and Econometrics of Networks (2013) conference, the Recent Advances in Set Identification (2013) conference, the 2nd European Meeting of Network Economics (2014), Heinz College (2015), the INET Econometrics of Networks Workshop at Cambridge University (2015), the Cowles Foundation (2015), Northwestern University (2015), the Chicago Fed (2015), the Econometric Society World Congress (2015), the USC Dornsife INET Conference on Networks (2015), the New York Fed (2016), the University of Pennsylvania (2016), Surrey (2016), Groningen (2016), the Tinbergen Institute (2016), Harvard (2016), Rutgers (2016), and Virginia (2016) for suggestions and comments. We are also grateful to A. Advani, J. Altonji, G. Chamberlain, C. Gualdani, P. Haile, T. Holden, and M. Shum for detailed comments. de Paula gratefully acknowledges financial support from the NSF through award SES-1123990, the ERC through Starting Grant 338187, and the Economic and Social Research Council through the ESRC Centre for Microdata Methods and Practice grant RES-589-28-0001.

[†]University College London, São Paulo School of Economics, C*e*MMAP, IFS and CEPR. E-mail: apaula@ucl.ac.uk

[‡]Lehigh University and NBER, E-mail: sethrs@lehigh.edu

[§]Harvard University, E-mail: elietamer@fas.harvard.edu

1 Introduction

This paper provides a framework for studying identification—what can be learned about parameters of interest from data—in strategic network formation models. The framework applies to complete information games with non-transferable payoffs in which individuals, given a particular utility function, form links with each other. Our objective is to learn about these payoffs from observed data on network linkages. In particular, we assume that observed networks are equilibrium networks and use pairwise stability, proposed in Jackson and Wolinsky (1996), as the solution concept.

A network formation model founded on a well-defined preference structure for all the players is helpful in determining how networks develop given a particular policy or incentive system. In a variety of applications, the architecture of networks is thought to influence final outcomes, and so it is important to understand how the networks themselves are formed. Empirically sound network formation models can also be helpful in describing why certain networks emerge and not others, for example tracking the effects of policies or frictions on the kinds and shapes of networks that arise.

The problem in analyzing such strategic models is that multiplicity of solutions and other computational difficulties are pervasive, especially when more than a handful of agents are involved. This paper proposes a framework for studying large networks, which achieves dimension reduction by aggregating individuals with similar network positions. We impose restrictions on the number of links that a person would choose to have, as well as on the cardinality of the observable characteristics. As a consequence, each individual's position in the network can be described as one of a finite number of possible local network structures. This greatly reduces the dimensionality of the problem. The limitation on the number of links, or bounded degree, is of course an important restriction. It makes this framework appropriate for applications where individuals are observed to have a relatively small number of connections, but not for others where it is important to have some nodes with many links (e.g., network "hubs").

Our approach for characterizing set-identified parameters bypasses the selection of a particular equilibrium (when many are possible) and directly exploits the economic predictions under pairwise stability. The key idea is that each individual can be classified as the central node (the "ego") of one of the finite possible local network structures, which we refer to as *network types*. The set of relevant types is determined by the preference specification, so for example a specification where only direct connections matter will lead to a different set of types than a specification where indirect connections matter as well. The link between the observed frequencies of these network types and their model predicted ones allows us to learn about preference parameters. A utility structure generates a set of payoff-relevant network types, and a parameterization of the network formation model predicts the measures or proportions of these types in the network. The inference question then reduces to collecting all utility parameters that can predict proportions of network types that match the proportions estimated in the data.

Developing this correspondence in a computationally feasible way represents the main contribution of this paper. Our main result provides necessary conditions for a set of parameters to generate a pairwise stable network with a distribution of network types that matches the observed distribution. For preference structures that satisfy an additional restriction we show that these conditions are sufficient as well, and so are able to characterize the identified set precisely. Then, to implement the approach, we show that the majority of the computation can be done through a series of quadratic programming problems.

In part because of the difficulties indicated above, the literature on the econometric analysis of strategic network formation models is small, but growing.[1] The payoff structures we analyze are related to those contemplated in, for example, Currarini, Jackson, and Pin (2009), Christakis, Fowler, Imbens, and Kalianaraman (2010), Sheng (2014), and (for a directed network) Mele (2017). The most similar papers to ours are those which also use pairwise stability and identify a set of parameters that could be consistent with the data in any equilibrium (Sheng 2014, Miyauchi 2016, Leung 2015).[2] Relative to these papers, our proposed method does not require certain restrictions on preferences (e.g., ruling out negative externalities or requiring a homophilous attribute), and it may be more computationally tractable. A number of other papers in this literature rely on dynamic meeting protocols for the formation of the network (Christakis, Fowler, Imbens, and Kalianaraman 2010, Mele 2017, Badev 2013).[3] Chandrasekhar and Jackson (2014) propose a different approach where the network is generated from overlapping sub-graphs. Also, some recent papers consider the estimation of dyadic link formation models (i.e., without link externalities) with a focus on disentangling homophily and node-specific heterogeneity (Graham 2017, Dzemski 2014, Charbonneau forthcoming).

[1]See Graham (2015) and de Paula (forthcoming) for recent surveys.

[2]Boucher and Mourifié (forthcoming) develop a method that is similar to Leung (2015) but bypasses the issue of multiplicity.

[3]Dynamic data on link formation are rarely available. As pointed out in Mele (2017), the meeting protocol therefore acts as an equilibrium selection mechanism. Additionally, Badev (2013) extends the model to include actions beyond link formation

Because matching models essentially aim at characterizing a bipartite graph, and hence a particular type of network, those models are also related to strategic network formation. There is a growing literature on the econometrics of matching (e.g., Choo and Siow (2006), Fox (2010a), Fox (2010b), Galichon and Salanie (2009), Echenique, Lee, and Shum (2010), Chiappori, Galichon, and Salanié (2016), Menzel (2015)). Our setting differs in substantive aspects however: indirect connections are payoff relevant, utility is non-transferable, and multiple equilibria are possible (in contrast to some papers in that literature). Also the concept of pairwise stability in matching games is related but not identical to the Jackson and Wolinsky (1996) definition, where only one link at a time is considered.

2 Network Formation Model

Our framework applies to complete information games that produce an undirected network. One example is a static game in which players simultaneously announce the set of other players they would like to be connected with, links form if they are mutually beneficial, and payoffs are received.[4] We use a continuum of players, $i \in N \equiv [0, \mu]$, where $\mu > 0$ is their total measure. This is unlike the typical setup in the empirical games literature, but the modeling choice is natural in our setting where there is a large number of players (Khan and Sun 2002). Each player has some predetermined characteristic(s) $X_i \in \mathcal{X}$ that is (are) observed by the econometrician, and player-pairs have a one-dimensional characteristic ϵ_{ij} that is not observed by the econometrician. Nature draws $\mathbf{X} = (X_i)_{i \in N}$ and $\epsilon = (\epsilon_{ij})_{(i,j) \in N^2}$, and these vectors are common knowledge to all players.

The network that results from players' actions is characterized by the adjacency mapping $G : N \times N \to \{0, 1\}$. This is a *continuous graph* as there is a continuum of nodes.[5] Such graphs (particularly refinements known as *graphings*, for limits of bounded degree graphs) are a recent development and are used as approximations for large graphs under a well-defined metric (for a review, see Lovasz (2012)). Hence we view the continuous graph model as a close approximation to a model with a large but finite number of players.[6]

Payoffs depend on the network configuration and covariates, and are denoted by $u_i(G, \mathbf{X}) = u(G, \mathbf{X}; \epsilon_i)$, where $\epsilon_i = (\epsilon_{ij})_{j \neq i}$. Our objective is to learn about the (parametric) utility func-

[4]Pairwise stability can also describe the rest point of a sequential meeting process (Jackson (2009)).

[5]Formally the graph consists of the adjacency mapping G and set of players N. However we typically refer to G as the "network."

[6]To be clear, we use the word "large" in its colloquial sense when referring to a network, not in the econometric sense that would imply the number of nodes is sent to infinity.

tions $u(G, \mathbf{X}; \cdot)$ using the data on G and $\mathbf{X}$.

To make the model tractable we rely on two main assumptions about the payoffs. We start with a restriction on network depth and total number of links.

Assumption 1. *Only connections up to distance D affect utility, and preferences are such that players will never choose more than a total of L links.*

The distance above refers to the length of the shortest path between two individuals, denoted $d(i, j; G)$. If $D = 1$ only direct connections are relevant (e.g., Currarini, Jackson, and Pin (2009)). When $D > 1$ indirect connections also matter, and here most specifications in the literature use $D = 2$ (e.g., "friends of friends"). The limit L denotes the maximum number of links an individual would have (i.e., utility would be infinitely negative if you have more than L links). This restricts our framework to networks with bounded degree distributions, where nodes have a relatively small number of links, rather than networks with approximately power-law degree distributions. Networks with such limited degree distributions are found in several social science contexts (e.g., close friendships) but not in others (e.g., Facebook).[7] Together, the restrictions on depth and degree in Assumption 1 make payoffs depend on a finite number of direct and indirect connections in the network. This is crucial for dimension reduction. For example with $D = 2$, there would be at most L direct alters and $L \times (L - 1)$ indirect alters that impact utility.

Our second assumption relates to the unobservable preference shocks, as well as the support of the observable characteristics. We assume that the preference shocks do not depend on the individual identities of the alters. Instead, there is one shock for each possible direct connection, combined with their possible predetermined characteristics. We further assume that the predetermined characteristics have finite support, and that the unobserved shocks are independent of these characteristics.

Assumption 2. *Individuals are endowed with $L \times |\mathcal{X}|$ preference shocks, denoted $\epsilon_l(x)$, $l = 1, \ldots, L$, $x \in \mathcal{X}$, which correspond to the possible direct connections and their characteristics. This vector of preference shocks is independent of X with a known distribution (possibly up to some finite dimensional parameter). In addition, the support of X is finite.*

[7] For example, such a limitation is seen in the National Longitudinal Study of Adolescent to Adult Health (also known as the "Add Health" study), a commonly used dataset on social networks. Individuals nominate up to five friends of each sex, and the number of reciprocated nominations is even smaller. The median number of such links is one, and less than five percent of individuals have more than three links to the same sex. Similarly, in data on networks among Indian villagers used in Jackson, Rodriguez-Barraquer, and Tan (2012), fewer than 1 per 1,000 respondents reached the caps of 5 or 8 nominations, respectively, for various kinds of social and financial relationships (footnote 37, p. 1879).

This assumption implies that if two potential alters have the same observables then the ego in question is indifferent between them. Similar (though not identical) assumptions about homogeneity in preferences have been made in models of large games (e.g., Kalai (2004), Menzel (2016)) and in some matching models (e.g., Choo and Siow (2006), Galichon and Salanie (2009)). This helps control the dimensionality of the problem, and it can be a natural restriction in settings with many agents where individual identities are unknown and irrelevant to the researcher. In addition, having a limited number of shocks allows the model to retain a positive fraction of isolated individuals in equilibrium even when the group under consideration is large.[8] This assumption could be weakened somewhat, by extending the vector of shocks to include ones for indirect connections up to distance D. In that case the number of shocks would be $|\mathcal{X}| \times L \sum_{d=1}^{D} (L-1)^{d-1}$ (for $L > 1$). Our theoretical results would be unchanged, but the computational burden would of course increase. Last, we note that Assumption 2 allows for correlation within the vector of preference shocks. This has economic relevance since it captures preference correlation among unobservables. However, generally allowing for unrestricted correlation might lead to exceedingly large identified sets.

Assumptions 1 and 2 together imply that there is a finite number of possible configurations of alters and their characteristics within the payoff-relevant distance from any given individual. Then, much as in discrete choice models for differentiated products, utility is determined from a fixed and finite list of possible outcomes. Our proposed inference strategy relies on this feature of the model to reduce the dimensionality of the problem from the universe of possible network configurations to the much smaller set of possible payoff-relevant local subnetworks, which we refer to as the network types in the model.[9]

Next we provide an example utility specification, which we will return to throughout the paper. Payoffs depend on the predetermined characteristics of the individual and her direct connections ($f(x_i, x_j) + \epsilon_{il}(x_j)$), any links among these direct connections (ω), and any further connections at distance 2 (ν). Such specifications are commonly used to model high school friendships, where the observable characteristics are race and perhaps other family attributes (e.g., Christakis, Fowler, Imbens, and Kalianaraman (2010), Mele (2017),

[8]Otherwise, if there were for example I.I.D. preference shocks for every potential connection in the network, the probability that a link is mutually beneficial with at least one other person increases as the size of the network increases (and hence the number of preference shocks increases).

[9]The use of subnetworks here is rather distinct from their use in Sheng (2014), although both are promising strategies for dimension reduction. We consider all possible subnetworks among individuals that are within some distance from a reference individual, where the distance is determined by the specification of preferences. Sheng (2014) considers subnetworks among arbitrary individuals, where the number of individuals in the subnetwork is chosen for computational tractability and is unrelated to the model. These approaches offer different advantages and should be viewed as complementary developments.

Goldsmith-Pinkham and Imbens (2013), Sheng (2014), Miyauchi (2016)). We use this context descriptively as well.[10] The utility function is expressed as follows:

$$
\begin{aligned}
u_i(G, \mathbf{X}) \equiv & \sum_{j \in N(i)} \Big(f(x_i, x_j) + \epsilon_{il(j)}(x_j) \Big) && \text{(direct connections)} \\
& + \Big| \bigcup_{j \in N(i)} N(j) - N(i) - \{i\} \Big| \nu && \text{(friends of friends)} \\
& + \sum_{j \in N(i)} \sum_{\substack{k \in N(i): \\ k > j}} G(j,k)\omega && \text{(mutual friends)} \\
& - \infty \cdot 1_{|N(i)|>L} && \text{(bounded degree)}
\end{aligned} \tag{1}
$$

where $N(i)$ denotes the neighbors of node i (i.e., $N(i) \equiv \{j : G(i,j) = 1\}$) and $|\cdot|$ gives the cardinality of a set. In the first line the index function $l(j)$ assigns neighbor j to the l^{th} link of node i and thereby assigns the preference shocks.[11] The second line takes the union of the neighbors of each friend but removes the friends themselves ($N(i)$) and the reference individual ($\{i\}$) to find the number of distinct friends of friends. The third line counts any links among the direct connections, and the fourth line ensures bounded degree in equilibrium. Like related specifications in the literature this uses a maximum depth of $D = 2$, and has additively separable shocks for direct connections from some known distribution.

As in most of the empirical games literature, we assume that observed choices correspond to equilibrium play. Our solution concept is pairwise stability (Jackson and Wolinsky 1996).

Definition 1 (Pairwise Stability)**.** *All links ij must be preferred by players i and j over not having the link, and all non-existing links must be damaging to at least one of the players:*

$$\forall i,j : G(i,j) = 1,\ u_i(G,\mathbf{X}) \geq u_i(G_{-ij},\mathbf{X}) \ \textit{and} \ \ u_j(G,\mathbf{X}) \geq u_j(G_{-ij},\mathbf{X}); \quad \textit{and} \tag{i}$$

$$\forall i,j : G(i,j) = 0,\ \textit{if} \quad u_i(G_{+ij},\mathbf{X}) > u_i(G,\mathbf{X}) \quad \textit{then} \quad u_j(G_{+ij},\mathbf{X}) < u_j(G,\mathbf{X}). \tag{ii}$$

[10] We believe the approach we put forward can accommodate many other applications. Our presentation is based on nontransferable utility (NTU). Applications elsewhere, for example in Industrial Organization (e.g., networks of buyers and suppliers or hospitals and insurers), may rely instead on transferable utility (TU). In that case, pairwise stability would require a positive surplus for observed links and a negative one for absent links, and preference classes, which we introduce later in the article, may need to rely on dyads rather than individuals. It may also be important to consider differences in numerosity whereby one side of the market (e.g., insurers) is substantially smaller than the other (e.g., hospitals). A careful examination these variations merits much more space and attention.

[11] The function $l(j)$ implicitly depends on G as well, because in different graphs the neighbor j could be assigned to different links. This is omitted to alleviate the complexity of the notation.

In the definition, G_{-ij} denotes the mapping $(k,l) \mapsto G_{-ij}(k,l) = G(k,l)$ if $(k,l) \neq (i,j)$ and $(k,l) \mapsto G_{-ij}(k,l) = 0$ if $(k,l) = (i,j)$. Analogously, G_{+ij} denotes the mapping $(k,l) \mapsto G_{+ij}(k,l) = G(k,l)$ if $(k,l) \neq (i,j)$ and $(k,l) \mapsto G_{+ij}(k,l) = 1$ if $(k,l) = (i,j)$. Other solution concepts exist, see Bloch and Jackson (2006) or Jackson (2009). As discussed in those references, an advantage of pairwise stability is that it incorporates the intuition that, in a social setting, agents are likely to communicate to form mutually desirable connections. This is not the case with Nash Equilibrium, where absent links can still be part of an equilibrium even though they would be mutually beneficial.

3 Preview of Our Results

Here we provide a preview of the approach using a very simple case of specification (1). In this example individuals have at most one link ($L = 1$, also $D = 1$), and there are two races: B (black) and W (white). This could describe a network of best friends, which consists only of isolates and linked pairs. There are no externalities from links in such a network, but it nevertheless illustrates the main features of our approach.

Outcomes can be expressed as ordered pairs, (x, y), for the individual's race (x) and the best friend's race (y, where $y = 0$ if no best friend). For example, (B, W) corresponds to a black individual with a white best friend. These pairs represent the *network types* in this model. (More generally, network types will involve a local adjacency matrix as well as a vector of node characteristics such as (B, W), but here the matrices are redundant.) Utility depends on an individual's network type, (x, y). The utility function (1) simplifies here to $u_i(x, y) = f_{xy} + \epsilon_i(y)$, where the f_{xy}, $x, y \in \{B, W\}$, are four parameters, and each individual is endowed with two preference shocks, $\epsilon_i(B)$ and $\epsilon_i(W)$, with some known distribution (up to a finite dimensional parameter). Our goal is to use data on the linkages in the network to learn about the parameters f_{xy}.

First, we collapse the global graph and node characteristics $(G, \mathbf{X})$ into the shares of individuals of each network type, or *type shares*. For example, suppose in a school with 500 students there were 50 black individuals with a white best friend. The share of type (B, W) is 0.1 (the share of type (W, B) is also 0.1, as they must balance). We will search for parameter values f_{xy} that can generate the observed type shares while satisfying necessary conditions for pairwise stability.

To do this, we start by classifying individuals based on which network types they would not reject (i.e., they would be happy with all the links). For example, depending on the preference shocks drawn, a black individual may prefer having a black best friend to being alone, but may not prefer having a white best friend to being alone: i.e., $f_{BB} + \epsilon_i(B) \geq 0$ and $f_{BW} + \epsilon_i(W) < 0$. Hence the network type (B, W) could not be an equilibrium outcome for this individual, but (B, B) could be. We refer to these sets of network types that individuals would not unilaterally deviate from as *preference classes*, generically denoted as H. For this individual, the preference class would be $H = \{(B, 0), (B, B)\}$. (Because there are no connections to be dropped from an isolated type, e.g. $(B, 0)$, every preference class contains one such type.) There are four possible preference classes for blacks in this example: $H_1 = \{(B, 0)\}$, i.e., prefers to be alone; $H_2 = \{(B, 0), (B, B)\}$, i.e., prefers a black best friend; $H_3 = \{(B, 0), (B, W)\}$, i.e., prefers a white best friend; $H_4 = \{(B, 0), (B, B), (B, W)\}$, i.e., prefers a best friend of either race. The preference classes for whites are similar, replacing the first race in each type with W.

Each preference class corresponds to some region in the space of the shocks, ϵ, that determines which network types would be "acceptable" to an individual with those shocks (i.e., an individual with shocks from the region for preference class H would not unilaterally deviate from any network type in H, but would deviate from any type not in H). Hence, given a distribution for the preference shocks and proposed values for the preference parameters, one can compute the probability that individuals fit into each preference class. In this example there are simple thresholds in $\epsilon(B)$ and $\epsilon(W)$ based on the parameters f_{xy} that yield these probabilities. With more elaborate models, the preference class probabilities can be computed, for example, via Monte Carlo integration (see Section 7 and Appendix D.7.1).

Given the preference class probabilities derived from a vector of structural parameters, we can then generate predicted type shares by allocating the individuals from each preference class to the possible network types. To do this we define *allocation parameters*, denoted $\alpha_H(\cdot)$, one for each type in each class, which designate the proportion of individuals allocated from preference class H to network type "$\cdot$". For example, the predicted share of blacks with a white best friend, type (B, W), is $\Pr(H_1|B)\ \alpha_{H_1}(B, W) + \Pr(H_2|B)\ \alpha_{H_2}(B, W) + \Pr(H_3|B)\ \alpha_{H_3}(B, W) + \Pr(H_4|B)\ \alpha_{H_4}(B, W)$ (multiplied by the proportion of blacks in the school to obtain the share among all students).

The key to our approach is to provide restrictions on the allocation parameters that need to be satisfied in order for the network to be pairwise stable. These restrictions correspond to necessary equilibrium restrictions and agreement with the data. First, individuals may only be allocated to network types contained in their preference classes. This restricts some of the allocation parameters to zero. Second, given any pair of network types that could feasibly add a link with each other (i.e., an isolated individual of either race in this example), the measure of individuals who would prefer to do so must be zero for at least one of the two types. Otherwise additional mutually beneficial links could be formed and the network would be unstable. Hence for any pair of types, the product of the measures of individuals of one type who would prefer to add links to individuals of the other type must be zero. This defines a quadratic objective function which in equilibrium has to be zero. Finally, the predicted proportions of types must match the observed proportions of types in the network, which defines a set of linear constraints.

4 Network Types and Preference Classes

Now we formalize and extend the concepts discussed in the preceding example. Our proposed identification strategy is built on the notion of pre-defined *network types*. These describe the local network structure around a given individual, along with the predetermined characteristics of each person (or node) in this local subnetwork. The size of the local subnetworks depends on the preference specification, specifically the parameters D and L that control the relevant depth in the network and maximum number of links. The predetermined characteristics are fixed attributes, such as sex and race, or predetermined behaviors (i.e., those which precede the formation of the network), for example the education levels of coworkers at a firm. Intuitively then, network types can be described in words, for example, "a female connected to two females and one male," "an unconnected low-income male," "a female connected to another female with two other friends," and so on.

More formally, a network type is characterized by a local adjacency matrix, A, and a vector of node characteristics, v. The matrix A describes the local subnetwork up to distance D from the reference individual, who is called the *ego* of the network type. It is symmetric and has one row and column for the ego and one for each possible alter up to distance D. The first row corresponds to the ego and indicates that individual's links with a 1 in the appropriate columns and 0, otherwise. The next L rows correspond to the possible direct alters, then the next $L(L-1)$ rows to the possible alters at distance 2, and so on. This gives a total of $1+L+L(L-1)+L(L-1)^2+\cdots+L(L-1)^{D-1} = 1+L\sum_{d=1}^{D}(L-1)^{d-1}$ rows.[12] The vector v contains the predetermined characteristics of the ego and the alters,

[12]This count simplifies to $1+L[(L-1)^D-1])/(L-2)$ lines when $L>2$, and is $1+2D$ when $L=2$ and simply 2 when $L=1$ (since links are reciprocated).

in the same order as the rows of A. The first element of v is the characteristic of the ego, denoted v_1. The subsequent elements are the characteristics of the possible alters, belonging to $\mathcal{X} \cup \{0\}$, where 0 denotes the absence of an alter in that position.[13] Thus we have:

Definition 2 (Network Type)**.** *Fix D, L, and $\mathcal{X}$. A network type t is characterized by $t = (A, v)$ where A is a square matrix of size $1 + L\sum_{d=1}^{D}(L-1)^{d-1}$ and v is a vector of same length as the number of rows in A. This matrix describes the local subnetwork that is utility-relevant for an individual of type t. The vector v contains the predetermined characteristics of this person and the alters in the local subnetwork. The complete enumeration of network types generated from a preference structure u and set of characteristics $\mathcal{X}$ is given by $\mathcal{T}$.*

As the definition says, the set of network types, $\mathcal{T}$, is determined completely by the preference structure and the set of predetermined characteristics. For example, if preferences are such that individuals have a taste for at most one friend (and there are no X's), then there would be only two types: "alone" and "connected." Each person in the network is one of these mutually exclusive and exhaustive types (i.e., each person can be identified as an ego of one particular network type), and we assume that the proportions of individuals of each network type can be consistently estimated from the observed data (see Section 6.2).

Below we use a particular example of specification (1) to illustrate the network types that arise from a given preference structure and to show their representation as (A, v).

Example 1. *Let preferences be given by* (1)*, with $D = 2$, $L = 2$, and $\mathcal{X} = \{B, W\}$. An individual has at most two direct alters and two indirect alters at distance 2.*[14] *The graphs of all the relevant network types under these preferences are visualized in Figure 1. The ego is represented with a square and the alters are circles. There are seven distinct graphs and a total of 62 different types, found by populating the nodes of these graphs with different combinations of the two races.*

Figure 2 shows the local adjacency matrix A and vector of predetermined characteristics v for two network types in this example. The first type is a triangle (graph 5 from Figure 1) where the reference individual (the ego) is black, with one black friend and one white friend who are also friends with each other. The first row of A indicates the ego's links to the two direct alters (in columns 2 and 3). The second and third rows correspond to the direct alters

[13] Typically there will be a collection of isomorphic matrix and vector pairs (A, v) that can characterize a given type t. The first row and column of A and the first element of v are reserved for the ego, but the remaining rows and columns of A and elements of v could be permuted and still express the same local subnetwork that is rooted in the ego. For computational convenience, we adopt a convention that singles out a representative characterization for each network type (see Appendix D.5.1.)

[14] We are grateful to an anonymous referee who suggested using an example with two links.

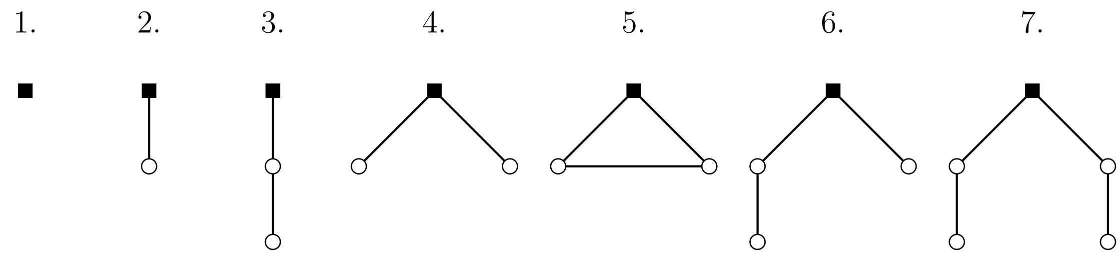


Figure 1: Graphs of network types in Example 1

First example type (graph 5):

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad v = \begin{pmatrix} B \\ B \\ W \\ 0 \\ 0 \end{pmatrix}$$

Second example type (graph 6):

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad v = \begin{pmatrix} B \\ B \\ W \\ B \\ 0 \end{pmatrix}$$

Figure 2: The matrix A and vector v for two network types in Example 1.

and indicate that they are linked to the ego and to each other. The second type is a tree (graph 6) where the ego is black with one black friend and one white friend, and the black friend has a further connection (i.e., a friend of a friend to the ego) who is also black. Here the direct alters are not linked to each other (elements (2,3) and (3,2) of A are 0), but the first direct alter (in row 2) is linked to an indirect alter represented in row and column 4.

Next, we note that under Assumptions 1 and 2 utility depends only on an individual's network type. Assumption 1 limits payoffs to depend on a subnetwork represented by A, and Assumption 2 restricts the unobserved shocks to depend on the characteristics listed in v, not the individual identities of the alters. Accordingly we rewrite the utility function as $u_i(G, \mathbf{X}) = u(A, v; \epsilon_i)$. The vector ϵ_i contains one preference shock for each potential direct connection and each possible characteristic of those connections (i.e., each element of $L \times |\mathcal{X}|$). In Example 1 there would be four shocks: $\epsilon_i = (\epsilon_{i1}(B), \epsilon_{i1}(W), \epsilon_{i2}(B), \epsilon_{i2}(W))$.[15] For the two example types shown in Figure 2, the utilities would be $f(B,B) + \epsilon_{i1}(B) + f(B,W) + \epsilon_{i2}(W) + \omega$ for the triangle with mutual friends and $f(B,B) + \epsilon_{i1}(B) + f(B,W) + \epsilon_{i2}(W) + \nu$ for the tree with a friend of a friend (the only difference is whether ω or ν appears).

In order to make predictions from the model, we will categorize individuals based on their preferences over network types. For this we define *preference classes*, the second important concept in our framework. These are sets of types that individuals would not unilaterally

[15] As noted earlier, this could be extended to include shocks for the indirect alters (friends of friends) as well. Then ϵ_i would have eight elements in total (four possible alters times two races), in this example.

deviate from, given their own preferences. In other words, individuals would not reject any of their links if assigned to one of the types in their preference class, but would reject a link if assigned to a type outside their preference class. Naturally, each preference class corresponds to a region in the support of ϵ: individuals with shocks in the same region (and with the same predetermined characteristics) would be content with the same set of types. The preference classes therefore partition the support of ϵ into regions that can rationalize different sets of network types, from an individual perspective.

Given the predetermined characteristics ($X_i = x$) and preference shocks (ϵ_i) of an individual, it is fairly straightforward to find his or her preference class. We compare the utility of each network type (where $v_1 = x$) against the utility that would be obtained by removing any of the links in that type, and retain all types that are not dominated by this. To express these comparisons, we define the matrix A_{-l} to be equal to the local adjacency matrix A but with the l^{th} link removed.[16] The preference class is then the set of types (A, v) with $v_1 = x$ such that: $u(A, v; \epsilon_i) \geq u(A_{-l}, v; \epsilon_i),\ \forall l = 1, \ldots, L$. We indicate how this works below.

Example (1 cont'd). *Suppose that the values of mutual friendships, ω, and friends of friends, ν, are positive. Consider an individual who is black ($X_i = B$), with preference shocks for black friends ($\epsilon_{i1}(B)$ and $\epsilon_{i2}(B)$) such that: $f_{BB} + \epsilon_{i1}(B) \geq 0$, $f_{BB} + \epsilon_{i2}(B) < 0$, $f_{BB} + \epsilon_{i2}(B) + \omega \geq 0$, and $f_{BB} + \epsilon_{i2}(B) + \nu \geq 0$. The first inequality above means that this person would prefer to have one black friend over being isolated. Hence the network type with graph 2 from Figure 1 and $v = (B, B, 0, 0, 0)'$ would be in this person's preference class. Also, because ν is positive, the network types with one friend and one friend of a friend—i.e., graph 3 and $v = (B, B, 0, B, 0)'$ or $v = (B, B, 0, W, 0)'$—would be in the preference class as well. The second inequality ($f_{BB} + \epsilon_{i2}(B) < 0$) means that i would not want to have a second black friend with no other connections. Hence the preference class would not include network types with graphs 4 or 6 from Figure 1, where the second friend is black. However the third and fourth inequalities above mean that the shock $\epsilon_{i2}(B)$ is large enough so that i would want to have a second black friend if there was a mutual friendship or an indirect connection (i.e., friend of friend) via this second friend. Hence the preference class would contain the network type with graph 5 where $v = (B, B, B, 0, 0)'$ and the types with graph 7 and $v = (B, B, B, v_4, v_5)$, where v_4 and v_5 are either B or W.*

The discussion above identifies seven types in individual i's preference class (and it rules

[16] In other words, set elements $(1, l+1)$ and $(l+1, 1)$ in A to zero (because alter number l appears in row $l+1$). There may remain nodes in the subnetwork represented by A_{-l} that are irrelevant to the resulting network type, because there is no path shorter than D between them and the ego once the l^{th} link is removed. All the entries in the rows and columns of A_{-l} corresponding to these nodes could be replaced with 0, as could the corresponding elements of v. Regardless, these nodes will have no impact on $u(A_{-l}, v; \epsilon_i)$.

out certain other types). Similar conditions on the shocks for white friends ($\epsilon_{i1}(W)$ and $\epsilon_{i2}(W)$) would determine which types with white friends, or a combination of black and white friends, are also in the preference class, thereby completing its construction.

To formalize the definition of preference classes, we introduce a mapping ψ that yields the preference class for a given value of (x, ϵ):

$$(x, \epsilon) \mapsto \psi(x, \epsilon) = H \subset \mathcal{T}, \tag{2}$$

where $H = \{(A, v) \in \mathcal{T} : v_1 = x \text{ and } u(A, v; \epsilon) \geq u(A_{-l}, v; \epsilon),\ \forall l = 1, \ldots, L\}$. The definition is then as follows.

Definition 3. *Given a preference structure u, let ψ be as defined in* (2). *A* **preference class** *$H \subset \mathcal{T}$ is a set of network types in the range of ψ.*

One important note is that every preference class includes an isolated type with the appropriate ego characteristics (e.g., in the example above, $A = 0_{5\times5}$ and $v = (B, 0, 0, 0, 0)'$). This is implicit from the definition of A_{-l}: there are no links to drop in the adjacency matrix for an isolated type, so $A_{-l} = A$. The presence of an isolated type in each preference class is important in order to allow for individuals who wind up isolated even if they desire connections. A second note is that preference classes must be obtainable under u from some vectors of characteristics and preference shocks. Hence the number of preference classes is smaller than the number of subsets of $\mathcal{T}$, and typically much smaller. That is because the preference structures in these models imply dependencies among the network types that can appear together in a preference class. This is seen in Example 1, where for example the positivity of ν implies that if a class includes a type with an unconnected friend (Figure 1, graph 2) it will also include types where that friend has a further friend (Figure 1, graph 3).

Finally, we can use the preference classes to generate predictions about the measures or proportions of individuals of each network type. First, we need the probability of each preference class. This is the probability that ϵ falls into the region corresponding to that class. We define these probabilities conditionally on the characteristics of the ego, as those characteristics are fixed within a given class. Accordingly the preference class probabilities are denoted as $P_{H|v_1} \equiv P(\epsilon : \psi(X, \epsilon) = H | X = v_1)$. These are direct functions of the utility specification and the distributions of the unobservables. Hence, given a parameterization of the model, these probabilities are known.

Then to generate the predictions, we specify how many individuals from each preference class are assigned to each network type. This uses *allocation parameters* defined as follows.

Definition 4. *An* **allocation parameter** $\alpha_H(t) \in [0,1]$ *gives the proportion of individuals with preferences in preference class* H *that are of network type* t*.*

The allocation parameters yield the measure of individuals of network type t as follows: $\mu_{v_1(t)} \sum_H P_{H|v_1(t)} \alpha_H(t)$, where $\mu_{v_1(t)}$ is the measure of individuals with characteristic $v_1(t)$ (the characteristic of the ego in type t). The proportion of individuals of network type t is this divided by the total measure μ. This provides the exact link between the data and the underlying preferences. The measures or proportions of individuals of each network type can be consistently estimated, and we will try to match these with predictions from the model.

5 Identification with Network Types

In this section, we show how to use the model to map the observed proportions of individuals of each network type (or more succinctly, the *type shares*) into restrictions on the preference parameters. We develop two general conditions on the allocation parameters that are necessary for pairwise stability, which can then be used to collect preference parameters that could be compatible with the observed type shares. If, using allocation parameters that satisfy these conditions, a vector of structural preference parameters cannot predict the observed type shares, then that vector is not compatible with the observed network. Otherwise, if such a prediction can be made, the vector is included in the recovered set.

The two conditions are as follows (their intuition is discussed after the theorem):

Condition 1 (Existing Links)**.** *All existing links are pairwise stable. For any type* t *and preference class* H*,* $t \notin H \implies \alpha_H(t) = 0$*.*

Condition 2 (Nonexisting Links Between Distant Individuals)**.** *There are no mutually beneficial links to add between individuals who are distant from each other in the network* ($d(i,j;G) > 2D$)*. For every pair of types* t, s *where the egos of both types have fewer than* L *links, and for the pair of types* $\bar{t}, \bar{s}$ *that would result if a link were added between two individuals of these types who are greater than* $2D$ *from each other,*

$$\left(\mu_{v_1(t)} \sum_{\tilde{H} \in \mathcal{H}} P_{\tilde{H}|v_1(t)} \alpha_{\tilde{H}}(t) 1_{\bar{t} \in \tilde{H}} \right) \cdot \left(\mu_{v_1(s)} \sum_{\check{H} \in \mathcal{H}} P_{\check{H}|v_1(s)} \alpha_{\check{H}}(s) 1_{\bar{s} \in \check{H}} \right) = 0.$$

The theorem below provides our general result on identification. It takes as given the predicted probabilities of the preference classes, $P_{\cdot|\cdot}$, which are yielded by a parameterization of the utility function. The theorem provides necessary conditions for a pairwise stable network with specified type shares (i.e., the observed shares) to exist given this distribution of preference classes in the population. To state the theorem we denote the vector of type shares as $\pi \equiv (\pi_t)_{t\in\mathcal{T}}$, where the element π_t is the proportion of individuals in the network who are of network type t. We maintain in this paper that our sample contains information on exactly these type shares.[17] For a given vector π and a distribution of preference classes $\{P_{H|v_1}\}$, we have the following result. (The proof is in Appendix A.)

Theorem 1. *Let the vector* $(\pi_t)_{t\in\mathcal{T}}$ *be known. Given a probability distribution of preference classes* $\{P_{H|v_1}\}$*, if there exists a pairwise stable network where the proportion of agents of type* t *is equal to* π_t *for each* $t \in \mathcal{T}$*, then there exists a vector of allocation parameters* α *satisfying Conditions 1 and 2 such that* π_t *is equal to* $\frac{1}{\mu}\sum_H \mu_{v_1(t)} P_{H|v_1(t)} \alpha_H(t)$ *for every* $t \in \mathcal{T}$*.*

This result can be extended to show that Conditions 1 and 2 are also sufficient for the existence of a pairwise stable network, under further restrictions on preferences. The main restriction is that the payoff from adding a link to a distant individual must be greater than that from adding a link to a nearby individual of the same type. Also a separability condition is required in the utility function. Because these restrictions are more specialized, the result on sufficiency is presented in Appendix B. Here we provide the intuition for how Conditions 1 and 2 translate the pairwise stability of links into conditions on the allocation parameters.

Condition 1 relates to expression (i) in the definition of pairwise stability. It is in fact equivalent because it requires individuals to be allocated only to network types in their preference classes, and those types must satisfy the same inequalities as in expression (i). Although Condition 1 treats individuals separately, this nevertheless implies that the inequalities hold for any pair of linked individuals. If instead Condition 1 is violated ($\alpha_H(t) > 0$ for some H and $t \notin H$), then there is some positive measure of individuals who would like to drop one of their links, and so the network cannot be pairwise stable.

Condition 2 relates to expression (ii) in the definition of pairwise stability. It establishes that there would be no further mutually beneficial links to add between individuals *who are at a distance greater than* $2D$ *from each other* in the network. This limitation to distances greater than $2D$ makes our conditions necessary but not sufficient for pairwise stability

[17] Our sampling approach to inference maintains that nodes are sampled randomly to collect information on types. This is common in the network statistics literature as it is common for networks to only be partially observed in practice (see, e.g., Chapter 5 in Kolaczyk (2009)).

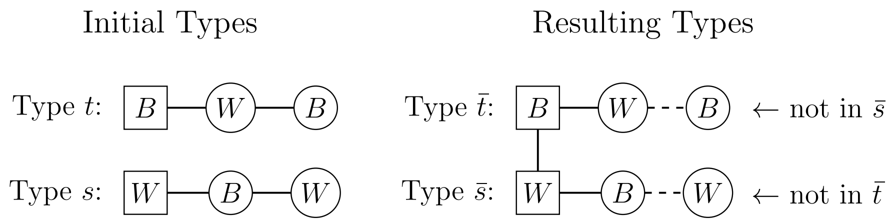


Figure 3: Example of a link added between individuals who are initially distant from each other. *Notes:* The ego of each type is represented with a square node. Dashed lines in the resulting types indicate connections to alters that appear in only one of the new network types, because they are beyond distance $D = 2$ from the ego of the other type.

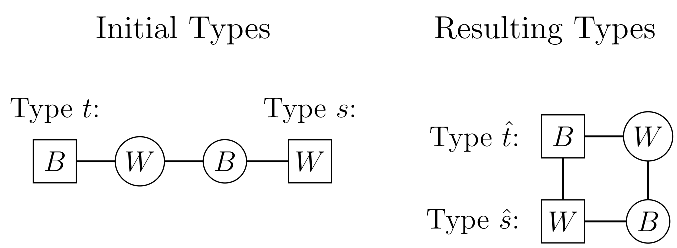


Figure 4: Example of a link added between individuals who are already close in the network

(except under the restrictions discussed in Appendix B). To understand Condition 2, note that there is one such equation for every pair of types (t, s), including pairs of the same type (t, t), where the egos of both types have fewer than L links. The other pair of types referred to in the condition, $(\bar{t}, \bar{s})$, would be obtained if a link were added between two individuals of types t and s who are greater than $2D$ from each other in the network. For a simple illustration, consider the types t and s shown in Figure 3, which follow Example 1 $(D = 2, L = 2)$. If a link were added between two individuals of these types, who were at a distance greater than $2D$, they would be transformed to the types $\bar{t}$ and $\bar{s}$ shown in the figure. This is because the local adjacency matrices for the resulting types would not capture any differences in the resulting structure of the (global) network, compared to the scenario in which the individuals were initially unconnected. These differences would involve loops of lengths greater than $2D+1$, which are not payoff-relevant and would not appear in the local adjacency matrices that extend only to distance D. Importantly, since the resulting types do not depend on the exact distance between the two individuals (so long as it is greater than $2D$), verifying Condition 2 does not require the global network G.

Nonexisting links between individuals who are $2D$ or less from each other are not considered in Condition 2 because different network types could result if a link were added. For example, if two individuals of the same types t and s from Figure 3 were initially connected at distance 3, as in Figure 4, adding a direct link would transform them to the types $\hat{t}$ and $\hat{s}$ shown in that figure. The utility of those types would potentially be different than for types $\bar{t}$ and $\bar{s}$ (e.g., there is only one friend of a friend in types $\hat{t}$ and $\hat{s}$). Assessing the stability of nonexisting links such as these would be more complex in our framework because, among other reasons, only certain network types can be located near each other (they must be able to have overlapping alters), and because the placement of individuals with different preference classes becomes important.

Next, note that with a positive measure of individuals of type s in the network, there are infinitely many individuals of type s who are beyond $2D$ from any one individual of type t (this is a consequence of the bounded degree assumption). Any of them could feasibly link with this individual of type t and transform her to type $\bar{t}$. So if this individual of type t prefers $\bar{t}$, and a positive measure of individuals of type s prefer $\bar{s}$ as well, the network is unstable. Accordingly Condition 2 requires that if a positive measure of type t individuals prefer $\bar{t}$ (i.e., $\alpha_{\tilde{H}}(t) > 0$ for some $\tilde{H}$ where $\bar{t} \in \tilde{H}$), there must be zero measure of type s individuals who prefer $\bar{s}$. Conversely, if a positive measure of type s individuals prefer $\bar{s}$, there must be zero measure of type t individuals who prefer $\bar{t}$. Notice that the expression $\mu_{v_1(t)} \sum_{\tilde{H} \in \mathcal{H}} P_{\tilde{H}|v_1(t)} \alpha_{\tilde{H}}(t) 1_{\bar{t} \in \tilde{H}}$ gives the total measure of type t individuals who prefer $\bar{t}$. Hence this or the analogous expression for the measure of type s individuals who prefer $\bar{s}$ must be zero. Because these measures cannot be negative, the condition that either one or the other measure must be zero is equivalent to requiring that their product be zero, as stated in Condition 2.

Finally, to complete the discussion, we show how these conditions relate to individual preferences in the context of Example 1.

Example (1 cont'd). *Consider the types t and s illustrated in Figure 3. Condition 1 requires that all individuals who are type t in the network have this type in their preference class. This means that $f_{BW} + \epsilon_1(W) + \nu \geq 0$, because the utility of type t is $f_{BW} + \epsilon_1(W) + \nu$, and the utility of the type obtained by removing the one link in this type is zero. Hence Condition 1 requires that all individuals who are type t prefer to have the link to their white friend (who has another friend). Similarly all individuals who are type s must have preferences in a class where $f_{WB} + \epsilon_1(B) + \nu \geq 0$, and so they prefer to have the link to their black friend (who has another friend).*

Condition 2 checks for individuals of type t who would prefer to be type $\bar{t}$, or type s who would prefer to be type $\bar{s}$. The expression inside the first parenthesis of the equation for these types gives the measure of individuals of type t ($\alpha_{\tilde{H}}(t)$ are allocated from each preference class

$\tilde{H}$) who also have type $\bar{t}$ in their preference classes ($1_{\bar{t}\in\tilde{H}}$). Such individuals would prefer to have a second white friend (who has another friend), because their preferences satisfy $f_{BW} + \nu + \epsilon_2(W) \geq 0$.[18] *The expression inside the second parenthesis gives the analogous measure of individuals of type s who also have $\bar{s}$ in their preference classes. One or the other of these measures must be zero. In other words, either none of the individuals of type t want a second white friend (who has another friend), or none of the individuals of type s want a second black friend (who has another friend), or both. If this condition is violated, there would be individuals of types t and s in the network who would prefer to be types $\bar{t}$ and $\bar{s}$ respectively. These individuals would prefer to add a link, and so the network would not be pairwise stable.*

Taken together, Conditions 1 and 2 restrict the preference classes that can be used to generate the predicted type shares for types t and s. Hence, in order to match the observed type shares, a given parameterization must place sufficient probability on the allowable preference classes for these types (i.e., those with type t but not $\bar{t}$ or those with type s but not $\bar{s}$).

The proof of Theorem 1 formalizes the preceding discussion by showing that, if expression (i) holds for all existing links and (ii) holds for all nonexisting links, then Conditions 1 and 2 must be satisfied. One note on this result is that it does not require the existence of equilibrium for every possible parameterization and realization of the variables (recall that nonexistence is possible under pairwise stability). If a particular parameterization cannot generate a pairwise stable network then there may be no vector of allocation parameters satisfying Conditions 1 and 2. In that case this parameterization would not be included in the identified set. If *no* parameterization can match the observed type shares while satisfying Conditions 1 and 2, then the identified set would be empty. We would conclude that the observation cannot be an equilibrium outcome under the model as specified, and so we might reject the model. Thus our framework can be used with models where nonexistence is possible, for example when links have negative externalities.

6 Implementation

We now describe how to use Theorem 1 to find values of the preference parameters that could be compatible with the observed network. First we show that the necessary conditions

[18]The utility of type $\bar{t}$ is $2f_{BW} + 2\nu + \epsilon_1(W) + \epsilon_2(W)$, and removing either link results in type t (note the symmetry of $\bar{t}$) with utility $f_{BW}+\nu+\epsilon_1(W)$. Hence type $\bar{t}$ is in a preference class when $f_{BW}+\nu+\epsilon_2(W) \geq 0$. Similarly type $\bar{s}$ is in a preference class when $f_{WB} + \nu + \epsilon_2(B) \geq 0$.

in the theorem can be verified using a quadratic programming (QP) problem.[19] Then we show how consistency of the estimators for the type shares from a single, large network can be obtained under a sampling approach to inference.

6.1 Formulation as Quadratic Programming Problem

Condition 2 provides a quadratic function of the allocation parameters that must equal zero in equilibrium. Using this to develop an objective function, a QP problem based on Theorem 1 can be defined as follows. The variables in the problem are those allocation parameters that are not set to zero by Condition 1: $\{\alpha_H(t) : t \in H\}$, or α for short. The objective function derived from Condition 2 is $\alpha^\top Q\alpha$, where the matrix Q is described in detail below. A set of linear constraints impose the requirement that the predicted type shares match the observed shares ($\frac{1}{\mu}\sum_H \mu_{v_1(t)} P_{H|v_1(t)}(\theta)\alpha_H(t) = \pi_t$). There are also adding-up and positivity constraints on the allocation parameters. The QP problem is thus

$$\min_{\{\alpha_H(t):t\in H\}} \alpha^\top Q\alpha \quad \text{subject to:} \tag{3}$$

$$\frac{1}{\mu}\sum_H \mu_{v_1(t)} P_{H|v_1(t)}(\theta)\alpha_H(t) = \pi_t, \forall t; \qquad \sum_{t\in H} \alpha_H(t) = 1, \forall H; \qquad \alpha_H(t) \geq 0.$$

As we establish further below, this problem has an optimal value of zero if and only if the conditions of Theorem 1 are satisfied. Therefore, given a vector of preference parameters (which produces a probability distribution of preference classes, $\{P_{H|v_1}(\theta)\}$), if a solution can be found yielding a value of zero, that parameter vector belongs in the recovered set.

The assembly of the programming problem is straightforward except for the objective matrix Q, which encodes Condition 2. This is a square matrix that has one row (and column) for each variable in the problem, so the entries of Q correspond to pairs of allocation parameters such as $\alpha_H(t)$ and $\alpha_G(s)$. The entries equal 1 for those pairs that could yield a positive value in the expression for Condition 2, and otherwise equal 0. Specifically, $Q_{[\alpha_H(t),\alpha_G(s)]} = 1_{\bar{t}\in H} \cdot 1_{\bar{s}\in G}$, where $\bar{t}$ and $\bar{s}$ are the types that would result if a link were added between two individuals of types t and s (as defined in the condition). This entry yields the term $(\alpha_H(t)\, 1_{\bar{t}\in H}) \cdot (\alpha_G(s)\, 1_{\bar{s}\in G})$ in the objective function $\alpha^\top Q\alpha$. Similarly, the expression in Condition 2 for this pair of allocation parameters includes the term

[19] A similar approach to "solving" for the identified set is seen in Honoré and Tamer (2006)'s use of a linear programming problem in a nonlinear panel data model. For a recent interesting example of the use of a quadratic programming problem in economics, see Kitamura and Stoye (2013).

$\left(\mu_{v_1(t)}\, P_{H|v_1(t)}\, \alpha_H(t)\, 1_{\bar{t}\in H}\right) \cdot \left(\mu_{v_1(s)}\, P_{G|v_1(s)}\, \alpha_G(s)\, 1_{\bar{s}\in G}\right)$. Hence, as long as $\mu_{v_1(\cdot)}$ and $P_{\cdot|\cdot}$ are strictly positive, the former can be used to assess whether the latter is nonzero.[20]

The example below illustrates the matrix Q in the simple model from Section 3.

Example 2. *Let preferences be given by* (1)*, with* $D = 1$*,* $L = 1$*, and* $\mathcal{X} = \{B, W\}$*. Here network types can be described using just the vectors of characteristics* $v = (v_1, v_2)$*. Preference classes can be enumerated as follows:* $H_1 = \{(B, 0)\}$*,* $H_2 = \{(B, 0), (B, B)\}$*,* $H_3 = \{(B, 0), (B, W)\}$*,* $H_4 = \{(B, 0), (B, B), (B, W)\}$*,* $H_5 = \{(W, 0)\}$*,* $H_6 = \{(W, 0), (W, W)\}$*,* $H_7 = \{(W, 0), (W, B)\}$*, and* $H_8 = \{(W, 0), (W, W), (W, B)\}$*. After excluding allocation parameters set to zero by Condition 1 there are 16 remaining:* $\alpha_1(B, 0)$*;* $\alpha_2(B, 0)$*,* $\alpha_2(B, B)$*;* $\alpha_3(B, 0)$*,* $\alpha_3(B, W)$*;* $\alpha_4(B, 0)$*,* $\alpha_4(B, B)$*,* $\alpha_4(B, W)$*;* $\alpha_5(W, 0)$*;* $\alpha_6(W, 0)$*,* $\alpha_6(W, W)$*;* $\alpha_7(W, 0)$*,* $\alpha_7(W, B)$*;* $\alpha_8(W, 0)$*,* $\alpha_8(W, W)$ *and* $\alpha_8(W, B)$ *(the subscripts correspond to the subscripts of the preference classes). The vector* α *collects these allocation parameters.*

The matrix Q *is* 16×16*, and its rows and columns correspond to the allocation parameters as listed in* α *(the matrix itself is shown in Appendix Figure 5). The matrix is symmetric, sparse, and indefinite. The first row corresponds to* $\alpha_1(B, 0)$*. All entries in that row are zero because the preference class associated with that allocation parameter,* $H_1 = \{(B, 0)\}$*, contains only the isolated type with a black ego. Hence* $1_{\bar{t}\in H_1} = 0$ *for any type* $\bar{t}$ *that could be obtained by adding a link. There are nonzero entries in six rows (and columns) of Q: those corresponding to allocation parameters* $\alpha_2(B, 0)$*,* $\alpha_3(B, 0)$*,* $\alpha_4(B, 0)$*,* $\alpha_6(W, 0)$*,* $\alpha_7(W, 0)$*, and* $\alpha_8(W, 0)$*. These parameters all indicate isolated individuals who would prefer to have a friend. (Because* $L = 1$*, only isolated individuals can add a link.) For example, in the row corresponding to* $\alpha_3(B, 0)$*, there are 1s in the columns corresponding to* $\alpha_7(W, 0)$ *and* $\alpha_8(W, 0)$*. The preference classes associated with these parameters are* $H_3 = \{(B, 0), (B, W)\}$*,* $H_7 = \{(W, 0), (W, B)\}$*, and* $H_8 = \{(W, 0), (W, W), (W, B)\}$*, respectively. To denote the types, let* $t = (B, 0)$ *(isolated black) and* $s = (W, 0)$ *(isolated white), and let* $\bar{t} = (B, W)$ *(black with white best friend) and* $\bar{s} = (W, B)$ *(white with black best friend), which are the the types that would result a link were added between two individuals of types* t *and* s*. Thus we have* $1_{\bar{t}\in H_3} \cdot 1_{\bar{s}\in H_7} = 1$ *and* $1_{\bar{t}\in H_3} \cdot 1_{\bar{s}\in H_8} = 1$*.*

The construction of Q in the example above is fairly simple because it is feasible to evaluate the expression $1_{\bar{t}\in H} \cdot 1_{\bar{s}\in G}$ for each entry individually. Given the corresponding pair of allocation parameters $\alpha_H(t)$ and $\alpha_G(s)$, the types t and s determine the types $\bar{t}$ and $\bar{s}$ that

[20] An expression identical to Condition 2 would be obtained if the entries of Q included the measures and probabilities; i.e., if $Q_{[\alpha_H(t),\alpha_G(s)]} = \left(\mu_{v_1(t)}\, P_{H|v_1(t)}\, 1_{\bar{t}\in H}\right) \cdot \left(\mu_{v_1(s)}\, P_{G|v_1(s)}\, 1_{\bar{s}\in G}\right)$. However having binary entries saves memory and potentially avoids recomputing Q at each putative parameter vector.

would result if a link were added, and it is then easy to check whether $\bar{t}$ and $\bar{s}$ are contained in H and G respectively. However for larger matrices it may be too burdensome to loop through the entries individually. Instead we suggest first constructing a preliminary matrix S, with the same dimensions and organization as Q, whose entries $S_{[\alpha_H(t),\alpha_G(s)]}$ are defined as $1_{\bar{t}\in H}$ rather than $1_{\bar{t}\in H} \cdot 1_{\bar{s}\in G}$. Conceptually the difference between S and Q is that the entries of S reflect only the preference class associated with the allocation parameter in the row, rather than those of both the row and the column. This makes it faster to construct S because operations can be applied row-by-row rather than element-by-element. The matrix Q then simply equals the Hadamard (i.e., entrywise) product of S with its transpose: $Q = S \circ S^\top$. (See Appendix D.7.2 for further details.)

With the objective matrix Q defined as above, we can establish the following result. (The proof appears in Appendix A.)

Theorem 2. *Given a probability distribution of preference classes* $\{P_{H|v_1}(\theta)\}$*, there exists a vector of allocation parameters* α *yielding type shares* $\{\pi_t\}$ *while satisfying Conditions 1 and 2 if and only if the optimal value of QP problem* (3) *is zero.*

This theorem provides a computational avenue to implement our approach. Because this pertains to identification, however, the population type shares are assumed to be known. In order to accommodate data from a finite sample, the QP problem must be modified to allow for some error in the estimated shares. The approach we take is to add slack variables that define fixed "bands" around the type shares, the width of which are a function of the sample size.[21] The modified QP problem then verifies whether a structural parameter vector can yield a prediction within these bands while satisfying Conditions 1 and 2.

Finally, we note that the objective function $\alpha^\top Q\alpha$ may not be convex because, while the matrix Q is symmetric, it may be indefinite as is the case in the example above. This rules out some standard QP solvers, but more general constrained nonlinear optimization routines can be used instead.[22] In the simulation exercises in Section 7, we find that the problem solves quickly using an active set algorithm in the program KNITRO. Importantly, because the optimal value is known (i.e., $\alpha^\top Q\alpha = 0$), it is trivial to ascertain that a global rather than local optimum has been reached. On the other hand, one must exercise caution so that positive local minima do not erroneously lead to dismissal of a parameter value.

[21]See Appendix D.6. Given a sampling distribution for the vector of estimated type shares, one could instead define the bands to contain a 95% confidence set. This would be a computationally efficient means to incorporate statistical uncertainty, as discussed in more general terms in Appendix C.

[22]Quadratic programs that are not convex are **NP**-hard. Alternative relaxation schemes are nonetheless available in the numerical optimisation literature to handle such cases.

6.2 Consistency of Type Shares

The parameter of interest in our setup is the vector that characterizes the payoff structure, θ, for a given large network. The main insight from Theorems 1 and 2 is that, *knowing the type shares* π, we have a mapping that yields the identified set for θ. We now briefly discuss inference and show how it is possible to obtain consistent estimates of π for this mapping.

Assuming that we do not observe the full network (and hence do not know π), we use a sampling approach to inference whereby we maintain that the observed data are a simple random sample from this full network that records the types only. The central question in this sampling approach is how close the sampled quantities are to the true quantities that can be obtained if we had access to the full network. Here the possibility of inverting from a sample to gain information on the population depends crucially on the sampling method employed. In our case, we maintain the assumption that we have a random sampling scheme where individuals are drawn independently (conditional on the realized network) and their types are recorded, which requires registering the features of their neighboring connections. The snowball sampling procedure defined in Goodman (1961), which starts with an initial random sample and interviews connections up to a specified distance, is an example of such a sampling scheme.[23] The only source of variation here is the random sampling, which then leads to statements about how close our estimate $\hat{\pi}$ is to π.[24]

Hence, while we do not know π, we can estimate it from appropriately sampled individuals in the network. As indicated in Section 2, the equilibrium graphing used in our theoretical analysis is an approximation to a large but nevertheless finite network, so here we consider a population network that is finite. Suppose we have data recording the network types t_i of a sample of n individuals. A sample analogue of the proportion of types is then $\hat{\pi}_t = \frac{1}{n}\sum_i 1[t_i = t]$. Accordingly, let $\hat{\pi} \equiv (\hat{\pi}_t)_{t\in\mathcal{T}}$, which is the vector of estimated type shares. Given that the population is finite, as n increases (i.e., as we are drawing a larger sample), it is then simple to see that $\hat{\pi}$ will converge to π. We state this result in the Proposition below.[25]

[23] The term "snowball sampling" has different uses. See Handcock and Gile (2011) for a helpful discussion.

[24] This relates to what is sometimes referred to as a "design-based" paradigm in statistics and goes back to at least Neyman (1934). The randomness here obtains from the probability ascribed by the survey scheme to the sampling of the various individuals in the network. In applications where the full network is observed (e.g., with administrative data) this approach would imply there is no statistical uncertainty.

[25] For a primer on the sampling approach to inference, see Cochran (2007). For sampling approaches with network data, see Kolaczyk (2009). Similar treatments using random sampling and relating finite and infinite populations are also given in Imbens and Rubin (2015) and references therein (see Chapter 6). Also, ARD data (Aggregated Relational Data) is an approach to collecting exactly sampled information (types) from a subset of nodes in a network. This ARD sampling approach to recovering full network properties from a

Proposition 1. *Under Random Sampling (on the underlying, realized network) and as sample size increases, we have:* $\hat{\pi} \xrightarrow{p} \pi$.

Appendix C contains the proof and elaborates on our approach in a simple example.[26]

7 Simulations

We now present two simulation exercises using examples of utility specification (1). The main purpose is to illustrate the performance of our approach, in terms of the parameter sets that are recovered and the computational burden that is involved. Additionally, the procedures described here and in Appendix D provide some guidance on further aspects of the implementation, such as how to generate the sets of network types and preference classes, and details of the construction of the objective matrix Q.

7.1 Model of Best Friendships

The first exercise uses the simple model of best friendships from Section 3, where $D = 1$, $L = 1$, and $\mathcal{X} = \{B, W\}$. Here network types can be represented using just the vector of characteristics $v = (x, y)$, where x is the race of the ego and y is the race of the alter (or 0 if no alter). The utility function simplifies to $u_i(x, y) = f_{xy} + \epsilon_i(y)$, with four preference parameters, f_{xy}, $x, y \in \{B, W\}$, and two preference shocks, $\epsilon_i(B)$ and $\epsilon_i(W)$.

The parameters are set at $f_{BB} = 0.4$, $f_{BW} = 0.2$, $f_{WB} = 0.15$, $f_{WW} = 0.5$, with population sizes $\mu_B = 1.0$, $\mu_W = 1.2$. The shocks have a uniform distribution (U$[-1, 0]$) so that the parameters can be interpreted as probabilities. We then use a microsimulation procedure to generate a single pairwise stable network with a large number of individuals (500 blacks and 600 whites).[27] From this network we extract the type shares, which are plotted in Figure 6. The figure also shows *all* possible type shares from pairwise stable networks under this parameterization, which demonstrates the full range of equilibria.[28]

random sample of nodes fits with our sampling approach (for the latest in econometrics on ARD work, see Breza, Chandrasekhar, McCormick, and Pan (2017).)

[26]There are alternative approaches to inference that take the repeated sampling approach whereby one observes one large network that is a draw from a population distribution of networks. There, large sample limit theory (where randomness is due to repeated sampling from this population distribution) is subtle due to the built-in correlations that result from interactions. For some recent advances, see Leung (2015) and Lee and Song (2017).

[27]This procedure is described in detail for the second excercise, in Appendix D.4.

[28]In this example, the QP problem simplifies so that it is trivial to verify whether a vector of preference

Given the type shares from the simulated network, we then search for parameter vectors where the QP problem can be minimized to zero. Figure 7 shows the resulting identified set. The values of the cross-race preference parameters (f_{BW} and f_{WB}) are unbounded from above, and together they display a Leontief pattern: if tastes on one side of the market constrain the number of cross-race linkages, tastes on the other side could be unbounded. By contrast, the identified ranges for the own-race preference parameters (f_{BB} and f_{WW}) are small and informative: 0.38–0.51 for blacks and 0.46–0.57 for whites. However the set would not provide conclusive evidence on the ranking of these parameters (i.e., $f_{WW} > f_{BB}$).

Finally, for this model, we are able to confirm the results obtained above and explore how the identified set changes with additional observations, using a simplification of the QP problem that can be derived in this particular case (see Appendix D.2). With the simplification it is trivial to verify that the optimal value is zero, and so we can evaluate a large grid of parameter vectors almost instantaneously. Figure A1 shows that the identified set matches what we obtained with the search procedure above. We then evaluate four other vectors of type shares, selected at random from the set of all equilibria shown in Figure 6. Qualitatively similar identified sets are recovered from each of these observations. (See Figure A2 for the vectors of type shares, and Figures A3 and A4 for the identified sets.) If we were to observe all four of the randomly selected networks, the identified set would be the intersection of the sets recovered with each network because the parameters must be able to predict all four vectors of type shares as equilibrium outcomes. As Figure 8 shows, the resulting identified set is quite precise. Only one parameter vector in the grid we evaluate is included ($f_{BB} = 0.40$, $f_{BW} = 0.20$, $f_{WB} = 0.16$, $f_{WW} = 0.50$). Since the grid uses intervals of 0.02, this means that the identified set spans less than 0.04 in each dimension. This indicates substantial identifying power from the observation of multiple large networks.

7.2 General Friendship Model

Next we consider a more complete version of specification (1), where $D = 2$ and $L = 3$. To motivate the latter, we note that fewer than five percent of the students in the Add Health study have more than three same-sex friends, based on reciprocated nominations. The predetermined characteristics are the same as before: black or white race, $\mathcal{X} = \{B, W\}$.

The preference parameters ($\theta \equiv (f_{BB}, f_{BW}, f_{WB}, f_{WW}, \nu, \omega)'$) are chosen to generate a

parameters and a vector of type shares are compatible with each other. This makes it easy to find *all* the equilibrium vectors of type shares for a given vector of structural parameters. See Appendix D.2 for details.

degree distribution that is similar to the observed distribution in Add Health, and the values of ν and ω also satisfy Assumption 3 so that the sharp identified set is obtained here (see Appendix B).[29] The six preference shocks are drawn independently from a standard normal distribution, but they are assigned to links in descending order (within alter characteristic). That way, the particular convention we use to select the representation, (A, v), for each network type does not impact their utility (see Appendix D.3).

To generate the data we simulate a number of finite networks with $n = 500$ individuals (100 blacks and 400 whites, reflecting $\mu_B/\mu_W = 1/4$), using a procedure described in Appendix D.4. Figure 9 plots the shares of certain types or combinations of types in these simulated networks, to illustrate the variation that can arise for a fixed vector of preference parameters but with different realizations of shocks and equilibria. One network, selected at random, serves as the observation we use to recover the identified set (indicated with triangles in Figure 9). Appendices D.6 to D.8 describe the specific procedures we then use to formulate and solve the QP problem (for a given parameter vector) and to search through the parameter space. One key point is that, to save memory and improve computational speed, we only consider network types that are either observed in the data or adjacent to an observed type (i.e., they can be reached via addition or deletion of one link). The search procedure uses Markov Chain Monte Carlo (MCMC) algorithms (Appendix D.8).

Projections of the identified set are shown in Figures 10 and 11. Figure 10 plots the parameters f_{xy} which govern the utility of direct connections. The identified range for f_{WB} appears to be unbounded from above, as in the previous example, while f_{BW} is bounded in both directions. Also, for both blacks and whites, we would not be able conclude that there is a preference for same-race over different-race friends (i.e., the fact that $f_{BB} > f_{BW}$ and $f_{WW} > f_{WB}$), as there are points in the identified set where the opposite holds. On the other hand, the values of same-race friendships (f_{BB} and f_{WW}) again have fairly tight ranges. Furthermore, if more than one network were observed, the identified set would be even smaller as seen in the previous example.[30] Figure 11 plots the identified values of ν (friends of friends) against the parameters for direct connections. The identified range of ν seems reasonably informative (it spans less than one standard deviation of the preference shocks), and its sign could be correctly inferred (the minimum identified value is 0.055).

[29] The parameter values are $(f_{BB}, f_{BW}, f_{WB}, f_{WW}) = (-0.9, -1.5, -1.7, -0.7)$, $\nu = 0.2$, and $\omega = 0.2$. Figure A5 in the Appendix shows the degree distributions. The average degree in the simulated networks matches the average degree for same-sex friendships in Add Health.

[30] In preliminary results from a simulation using two observations, the identified range for f_{WB} is bounded and is roughly similar in size to the identified range for f_{BW}, for example.

The parameter ω (mutual friends) is not shown. The data do not provide any information on its value, mainly because there are no mutual friendships in the observed network. The identified range for ω is consequently unbounded from below, and the upper bound ($\omega \leq \frac{3}{2}\nu$) follows from Assumption 3 (see Appendix B).

Importantly, the computational burden involved in this exercise indicates that our approach is feasible for empirically realistic models. The average time required to evaluate a candidate vector of structural parameters was just less than 30 seconds (see Appendix D.8). Furthermore, the MCMC search process is trivial to parallelize by running multiple chains simultaneously, so richer models with potentially longer evaluation times (say 5 to 15 minutes) would remain tractable. Most of the computational burden (80% of the total compute time) comes from solving the QP problem, which also suggests that additional gains in performance may be possible with further advances in solution algorithms.

8 Conclusion

We conclude with a brief discussion about empirical settings where our framework might apply, and where it might not. The main consideration is whether the assumption of bounded degree is reasonable in a particular context. The maximum degree and the degree distribution are of course observable in a given dataset, and depending on the sampling scheme (and the model) one can determine whether a specific bound is reasonable to impose. For example in the social and financial networks studied in Jackson, Rodriguez-Barraquer, and Tan (2012), a negligible portion of the sample reached the caps of 5 or 8 nominations for each type of relationship. This suggests that the bounded degree assumption may be reasonable in analyses of informal insurance as well as close friendships. On the other hand bounded degree would not apply in online social networks. Nor would it apply in certain markets for intermediate goods or financial networks where firms with many connections (i.e., "hubs") are observed. Economically, the difference between settings with bounded degree distributions vs. approximately power law distributions may relate to the cost of links. If there are substantial fixed costs, relative to a finite budget of time for example, individuals would be limited in the number of links they can maintain.

References

Badev, A. (2013): "Discrete Games in Endogenous Networks: Theory and Policy," University of Pennsylvania Working Paper.

Bloch, F., and M. Jackson (2006): "Definitions of Equilibrium in Network Formation Games," *International Journal of Game Theory*, 34, 305–318.

Boucher, V., and Mourifié (forthcoming): "My Friend Far Far Away: Asymptotic Properties of Pairwise Stable Networks," *Econometrics Journal*.

Breza, E., A. G. Chandrasekhar, T. H. McCormick, and M. Pan (2017): "Using Aggregated Relational Data to feasibly identify network structure without network data," Discussion paper, National Bureau of Economic Research.

Byrd, R. H., N. I. Gould, J. Nocedal, and R. A. Waltz (2003): "An algorithm for nonlinear optimization using linear programming and equality constrained subproblems," *Mathematical Programming*, 100(1), 27–48.

Chandrasekhar, A. G., and M. O. Jackson (2014): "A Network Formation Model Based on Subgraphs," Working Paper 20276, NBER.

Charbonneau, K. (forthcoming): "Multiple fixed effects in binary response panel data models," *Econometrics Journal*.

Chiappori, P.-A., A. Galichon, and B. Salanié (2016): "On Human Capital and Team Stability," Columbia University Working Paper.

Choo, E., and A. Siow (2006): "Who Marries Whom and Why," *Journal of Political Economy*, 114(1), 175–201.

Christakis, N., J. Fowler, G. Imbens, and K. Kalianaraman (2010): "An Empirical Model of Strategic Network Formation," Harvard University Working Paper.

Cochran, W. G. (2007): *Sampling techniques*. John Wiley & Sons.

Currarini, S., M. Jackson, and P. Pin (2009): "An Economic Model of Friendship: Homophily, Minorities and Segregation," *Econometrica*, 77(4), 1003–1045.

de Paula, A. (forthcoming): "Econometrics of network models," in *Advances in Economics and Econometrics: Theory and Applications, Eleventh World Congress*. Cambridge University Press.

Dzemski, A. (2014): "An Empirical Model of Dyadic Link Formation in a Network with Unobserved Heterogeneity," University of Manheim Working Paper.

Echenique, F., S. Lee, and M. Shum (2010): "Aggregate Matchings," Caltech Working Paper.

Fox, J. (2010a): "Estimating Matching Games with Transfers," University of Michigan Working Paper.

Fox, J. T. (2010b): "Identification in Matching Games," *Quantitative Economics*, 1(2), 203–254.

Galichon, A., and B. Salanie (2009): "Matching with Trade-offs: Revealed Preferences over Competing Characteristics," Ecole Polytechnique Working Paper.

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

Goodman, L. A. (1961): "Snowball Sampling," *The Annals of Mathematical Statistics*, 32(1), 148–170.

Graham, B. S. (2015): "Methods of Identification in Social Networks," *Annual Review of Economics*, 7(1), 465–485.

——— (2017): "An Econometric Model of Network Formation with Degree Heterogeneity," *Econometrica*, 85(4).

Handcock, M. S., and K. J. Gile (2011): "Comment: On the Concept of Snowball Sampling," *Sociological Methodology*, 41(1), 367–371.

Honoré, B. E., and E. Tamer (2006): "Bounds on Parameters in Panel Dynamic Discrete Choice Models," *Econometrica*, 74(3), 611–629.

Imbens, G. W., and D. B. Rubin (2015): *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Jackson, M. (2009): *Social and Economic Networks*. Princeton University Press.

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

Jackson, M. O., T. Rodriguez-Barraquer, and X. Tan (2012): "Social Capital and Social Quilts: Network Patterns of Favor Exchange," *American Economic Review*, 102(5), 1857–97.

Kalai, E. (2004): "Large Robust Games," *Econometrica*, 72(6), 1631–1665.

Khan, M. A., and Y. Sun (2002): "Non-Cooperative Games with Many Players," vol. 3 of *Handbook of Game Theory with Economic Applications*, chap. 46, pp. 1761–1808. Elsevier.

Kitamura, Y., and J. Stoye (2013): "Nonparametric Analysis of Random Utility Models," CeMMAP CWP36/13.

Kolaczyk, E. (2009): *Statistical Analysis of Network Data*. Springer-Verlag.

Lee, J. H., and K. Song (2017): "Stable Limit Theorems for Empirical Processes under Conditional Neighborhood Dependence," UBC Working Paper.

Leung, M. (2015): "A Random-Field Approach to Inference in Large Models of Network Formation," Stanford University Working Paper.

Lovasz, L. (2012): *Large Networks and Graph Limits*. American Mathematical Society.

Mele, A. (2017): "A Structural Model of Dense Network Formation," *Econometrica*, 85(3), 825–850.

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

——— (2016): "Inference for Games with Many Players," *Review of Economic Studies*, 83, 306–337.

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

Neyman, J. (1934): "On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection," *Journal of the Royal Statistical Society*, 97, 558–625.

Sheng, S. (2014): "A Structural Econometric Analysis of Network Formation Games," UCLA Working Paper.

van der Vaart, A. W. (1998): *Asymptotic Statistics*. Cambridge: Cambridge University Press.

A Proofs of Theorems 1 and 2

A.1 Theorem 1

Proof. Given a pairwise stable network G, predetermined characteristics $\mathbf{X}$, and preference shocks ϵ (for all the individuals in the network), partition individuals by their preference class H and define $\alpha_H(t)$ as the fraction of individuals in this pairwise stable network in preference class H who are of network type t. This allocation yields the observed proportions of network types; i.e., $(\mu_{v_1(t)}/\mu)\sum_H P_{H|v_1(t)}\alpha_H(t) = \pi_t$.

If the network G is pairwise stable, then all existing links satisfy expression (i) in Definition 1. Hence for any individual i, whose network type is characterized as (A, v), we have $u(A, v; \epsilon_i) \geq u(A_{-l}, v; \epsilon_i)$, $\forall l = 1, \ldots, L$. By definition this type is in the individual's preference class. Hence $\alpha_H(t) > 0$ only if $t \in H$, and so Condition 1 is satisfied.

To show that Condition 2 is satisfied, consider an arbitrary pair of types (t, s) where both types have fewer than the maximum L links. Let $(\bar{t}, \bar{s})$ be the pair of types that would result if a link were added between two individuals of types t and s who are greater than $2D$ from each other. Suppose that the measure of individuals of type t who prefer $\bar{t}$ is positive: $\mu_{v_1(t)}\sum_{\tilde{H}\in\mathcal{H}} P_{\tilde{H}|v_1(t)}\alpha_{\tilde{H}}(t)1_{\bar{t}\in\tilde{H}} > 0$. Hence there is an infinite number of individuals of type t who would prefer to add links to some individuals of type s that are greater than $2D$ away from them. These individuals of type t would indeed prefer to add links to *any* individuals of type s that are greater than $2D$ away, because adding a link to any distant individual of type s would yield the same utility for a given individual of type t. Also, because degree is bounded, only a finite number of individuals of type t can be within distance $2D$ from any given individual of type s. Hence every individual of type s faces individuals of type t who are greater than $2D$ away and who want to add a link. Because all nonexisting links satisfy expression (ii) in a pairwise stable network, the utility of type $\bar{s}$ must therefore be lower than the utility of type s for *all* individuals of type s. Because type s can be obtained by dropping one of the links in type $\bar{s}$, this means that $\bar{s}$ is not contained in the preference class of any individuals of type s in the network. Therefore the measure $\mu_{v_1(s)}\sum_{\check{H}\in\mathcal{H}} P_{\check{H}|v_1(s)}\alpha_{\check{H}}(s)1_{\bar{s}\in\check{H}}$ is zero. Thus at least one of the measures expressed in the equation for types (t, s) in Condition 2 must be zero, which gives us the condition. □

A.2 Theorem 2

Proof. Condition 2 is satisfied if and only if the objective function is equal to zero. This is because, as long as $\mu_{v_1(\cdot)}$ and $P_{\cdot|\cdot}$ are strictly positive,

$$\mu_{v_1(t)}\mu_{v_1(s)} \sum_{\tilde{H}\in\mathcal{H}} \sum_{\check{H}\in\mathcal{H}} P_{\tilde{H}|v_1(t)} P_{\check{H}|v_1(s)} \alpha_{\tilde{H}}(t)\alpha_{\check{H}}(s) 1_{\bar{t}\in\tilde{H}} 1_{\bar{s}\in\check{H}} = 0$$
$$\Leftrightarrow$$
$$\sum_{\tilde{H}\in\mathcal{H}} \sum_{\check{H}\in\mathcal{H}} \alpha_{\tilde{H}}(t)\alpha_{\check{H}}(s) 1_{\bar{t}\in\tilde{H}} 1_{\bar{s}\in\check{H}} = 0.$$

The first set of constraints is to match the observed proportions of network types (π_t). The second and third sets of constraints simply require that allocations from a given preference class add up to one and are non-negative. Finally, Condition 1 is encoded by the fact that allocation parameters are only defined for the types in each preference class (i.e., the variables in the problem are $\{\alpha_H(t) : t \in H\}$, not all $\{\alpha_H(t)\}$). Hence from each preference class there are no allocations made to types not in that preference class. □

B Sufficiency and Sharp Identification

Conditions 1 and 2 are sufficient (as well as necessary) for the existence of a pairwise stable network under further restrictions on preferences, which may be reasonable in some applications. The main restriction is that the marginal utility from adding a link to a distant individual ($> 2D$) must be greater than that from adding a link to a nearby individual ($\leq 2D$) of the same type. Also a form of separability is required in the utility function, as defined below. Then, because Conditions 1 and 2 are necessary *and* sufficient for equilibrium, we can use them to recover exactly the set of preference parameters that are compatible with the observed type shares (i.e., the sharp identified set).

To state the main restriction, we need to denote the type(s) that an individual of type t could become if he or she were to add a link to an individual of type s who is within distance $2D$. So, abstractly, we define the correspondence $\Psi(t, s)$ to collect all such types.[31] For example, in Figure 4, $\hat{t} \in \Psi(t, s)$ and $\hat{s} \in \Psi(s, t)$ (note the arguments are reversed). In general, these sets could be constructed via enumeration: first check each alter node in type t to see whether an individual of type s could be located there, then determine what type

[31] There can be multiple types in $\Psi(t, s)$ if it is feasible to have alters of type s located at multiple nodes in the subnetwork defined by type t.

the individual of type t would become if a link were added to an individual of type s at that location. To define the separability requirement, we refer to separate components in the network type *when the ego is removed.* These components consist of subsets of alters that are connected with each other (i.e., paths exist among them) but not with other alters in the network type, if the ego is removed. We require additive separability across such subsets of alters. Bringing these together, the assumption is as follows.

Assumption 3. *For an individual of type t, let $\bar{t}$ be the network type that would be obtained by linking to an individual of type s who is at a distance greater than $2D$, and let $\Psi(t,s)$ collect the type(s) that could be obtained by linking to an individual of type s who is at a distance of $2D$ or less. Assume the following.*
(a) $u(\bar{t};\epsilon) \geq u(\hat{t};\epsilon), \forall \hat{t} \in \Psi(t,s)$
(b) For any type t, the utility function is additively separable across sets of edges and vertices that belong to different components of the graph obtained when the ego is removed from the network type.

The first part of this assumption is plainly at odds with the clustering observed in many networks, because it does not encourage links among individuals who already share connections. Nevertheless it is valid for certain preference structures that have received attention in the literature. It is trivially satisfied in models where only direct connections matter (e.g., Currarini, Jackson, and Pin (2009)), because the resulting types yield the same utility (i.e., $u(\bar{t};\epsilon) = u(\hat{t};\epsilon)$). It is also satisfied in some models with indirect connections, such as the connections model in Jackson and Wolinsky (1996). The key point is that Assumption 3(a) is more a statement about marginal utility than about total utility.[32] It is possible for network types with more nearby alters to provide greater total utility, while adding links to more distant nodes yields greater marginal utility. This holds in the connections model because the marginal utility of adding a link to some node is reduced by the current value of the indirect connection to that node, which decreases in distance.[33] More generally it seems that this would be the case when one's closeness (or distance) to other nodes in the network matters more than the clustering among one's alters. In our example utility specification (1), which values both closeness and clustering in some sense, Assumption 3(a) is satisfied in a

[32]Note that Assumption 3(a) is equivalent to $u(\bar{t};\epsilon) - u(t;\epsilon) \geq u(\hat{t};\epsilon) - u(t;\epsilon)$.

[33]In the connections model adding a link to a node currently at distance d yields $w_{ij}(\delta^1 - \delta^d) - c_{ij}$ (where $w_{ij} \geq 0$ is the benefit of being connected to node j, $0 < \delta < 1$ discounts that benefit by the distance between i and j, and c_{ij} is the cost of a direct link with node j). The marginal utility of adding link ij also depends on changes in the distances to other nodes reached via node j, and those marginal benefits are similarly reduced by the current distance to j.

subset of the parameter space where the value of mutual friendships (ω) is bounded relative to the value of friends of friends (ν). (This is shown in Corollary 1 after the theorem.)

The second part of Assumption 3 makes the utility function additive across links going to separate subsets of alters (i.e., in separate components of the graph without the ego, as defined above). For example, the utility of type $\bar{t}$ in Figure 3 must be the sum of the payoffs obtained from having each of the two links. (Removing the ego from this type leaves two components, each consisting of a W linked to a B, so the assumption is binding.) On the other hand, the alters in type $\hat{t}$ in Figure 4 remain connected without the ego, so the utility of this type does not need to be separable across the ego's two links. Utility specification (1) indeed satisfies Assumption 3(b) for any values of the parameters, while allowing for important complementarities across links such as with mutual friendships (see Corollary 1). However the assumption is restrictive in general, for example ruling out convex costs in the number of links. (Having convex costs might be natural in many applications, but in most specifications in the econometric literature on networks they are not included.)

With this assumption, we then state the result as follows.[34]

Theorem 3. *For models satisfying Assumption 3, given a probability distribution of preference classes $\{P_{H|v_1}\}$, there exists a network that is pairwise stable (except possibly for a set of agents with zero measure) where the proportion of agents of type t is equal to π_t for each $t \in \mathcal{T}$ if and only if there exists a vector of allocation parameters α satisfying Conditions 1 and 2 such that $\pi_t = \frac{1}{\mu}\sum_H \mu_{v_1(t)} P_{H|v_1(t)} \alpha_H(t)$ for every $t \in \mathcal{T}$.*

The necessity of Conditions 1 and 2 was established by Theorem 1. The proof of their sufficiency below uses a contrapositive argument. It shows that if there is no pairwise stable network with the given type shares, then allocation parameters cannot be found satisfying the two conditions. The key point is that if positive measures of nearby individuals mutually desire to add links, then under Assumption 3(a) these individuals must also desire to add links with distant individuals of the same types, which violates Condition 2.

Proof. To start, fix a vector of preference parameters and, hence, a distribution of preference classes, as well as the observed proportions of network types. Suppose that under these preferences, any network with these type shares is unstable. Therefore, for any such network there must be a positive measure of pairs of individuals for whom the presence or absence

[34] Because Conditions 1 and 2 assess the measures of sets of agents, they cannot guarantee pairwise stability for sets with zero measure. This issue does not arise in Theorem 1 because it does not affect the necessity of the conditions.

of a link between them is unstable. To translate this into our conditions, first note that for any network among a set of players N, there is a unique vector of allocation parameters that expresses the allocation of the individuals from each preference class to each network type. This is because each individual is associated with one and only one preference class, and one and only one network type.

First we consider existing links ($G(i,j) = 1$). If there is a positive measure of pairs of individuals who are linked but one or both of them would prefer to drop the link, then there must be some preference class H where a positive measure of individuals with preferences in this class are some network type that is not in H. Therefore $\alpha_H(t) > 0$ for some $t \notin H$ and so Condition 1 would be violated.

If all existing links are stable, there must be nonexisting links ($G(i,j) = 0$) that are unstable (i.e., both i and j would prefer to add the link). We first consider nonexisting links between individuals who are distant from each other in the network (i.e., $d(i,j;G) > 2D$). If there is a positive measure of such pairs of individuals who would prefer to be linked with each other, then there is at least one pair of network types (t,s) such that positive measures of individuals of these two types would prefer to add links with each other. A link between two individuals of types t and s who are distant from each other would transform them to types $\bar{t}$ and $\bar{s}$ respectively, and this tuple of types, (t,s) and $(\bar{t},\bar{s})$, pertains to one of the equations in Condition 2.

We now use Assumption 3(b) to show that $\bar{t}$ and $\bar{s}$ are in the respective preference classes of these individuals of types t and s, which leads to a violation of Condition 2. Consider $\bar{t}$. This type contains the same subnetwork of alters as type t, plus an additional subnetwork from the link to an individual of type s. If the ego is removed from type $\bar{t}$, the additional subnetwork is a separate component in the graph (there are no other paths from the ego in $\bar{t}$ to the alter formerly of type s, because that alter was previously beyond $2D$). The utility of type $\bar{t}$ is therefore additively separable between the subnetwork from type t and the additional subnetwork. Denote the links in type t as $l = 1 \ldots l_t$, the additional link in type $\bar{t}$ as $l_t + 1$, and the representations of these types as (A_t, v_t) and $(A_{\bar{t}}, v_{\bar{t}})$ respectively. (For simplicity, suppose that the alters from type t appear in the same rows of $A_{\bar{t}}$ as they do in A_t.) Given the additive separability, if type t is in an individual's preference class (i.e., $u(A_t, v_t; \epsilon) \geq u(A_{t,-l}, v_t; \epsilon)$, $l = 1 \ldots l_t$), then the corresponding links in type $\bar{t}$ are preferred as well (i.e., $u(A_{\bar{t}}, v_{\bar{t}}; \epsilon) \geq u(A_{\bar{t},-l}, v_{\bar{t}}; \epsilon)$, $l = 1 \ldots l_t$). This indeed holds for the relevant individuals of type t, because here we are supposing that all existing links are stable (otherwise Condition 1 is violated). These individuals also prefer type $\bar{t}$ over type t, so the additional link in $\bar{t}$ is preferred as well (i.e., $u(A_{\bar{t}}, v_{\bar{t}}; \epsilon) \geq u(A_{\bar{t},-(l_t+1)}, v_{\bar{t}}; \epsilon)$). Therefore all the links in type $\bar{t}$ satisfy the preference inequalities needed for this type to be in the preference classes of these individuals. Hence the expression $\mu_{v_1(t)} \sum_{\tilde{H} \in \mathcal{H}} P_{\tilde{H}|v_1(t)} \alpha_{\tilde{H}}(t) 1_{\bar{t} \in \tilde{H}}$ is strictly positive. The same argument applies for type $\bar{s}$ and the relevant individuals of type s. Hence the expression $\mu_{v_1(s)} \sum_{\check{H} \in \mathcal{H}} P_{\check{H}|v_1(s)} \alpha_{\check{H}}(s) 1_{\bar{s} \in \check{H}}$ is strictly positive as well. Therefore the product of the measures given by these expressions is strictly positive, which violates Condition 2.

Last, we consider nonexisting links between individuals who are $2D$ or less from each other in the network. If there is a positive measure of such pairs of individuals who would prefer to be directly linked, then there is some pair of network types (t, s) and some distance $d \leq 2D$ such that positive measures of individuals of these two types who are at distance d from each other would prefer to add links with each other. A link between two such individuals would transform them to some types $\hat{t} \in \Psi(t, s)$ and $\hat{s} \in \Psi(s, t)$ respectively. These individuals prefer $\hat{t}$ over t and $\hat{s}$ over s, respectively. By Assumption 3(a), they weakly prefer $\bar{t}$ over $\hat{t}$ and $\bar{s}$ over $\hat{s}$, respectively (where, as before, the types $\bar{t}$ and $\bar{s}$ would result from a link being added between distant individuals of types t and s). Hence they prefer $\bar{t}$ over t and $\bar{s}$ over s, respectively. Then the same argument applies as for nonexisting links between distant individuals. For these individuals, all the links in type t (respectively, s) are preferred, as is the additional link in type $\bar{t}$ ($\bar{s}$), hence $\bar{t}$ ($\bar{s}$) is in their preference classes. Therefore, as above, the expressions $\mu_{v_1(t)} \sum_{\tilde{H} \in \mathcal{H}} P_{\tilde{H}|v_1(t)} \alpha_{\tilde{H}}(t) 1_{\bar{t} \in \tilde{H}}$ and $\mu_{v_1(s)} \sum_{\check{H} \in \mathcal{H}} P_{\check{H}|v_1(s)} \alpha_{\check{H}}(s) 1_{\bar{s} \in \check{H}}$ are strictly positive, and so Condition 2 would be violated.

□

Thus in cases where Assumption 3 is appropriate, Conditions 1 and 2 can be used to characterize the identified set precisely. We next show how this applies to specification (1).

Corollary 1. *In utility specification* (1)*, if* $\nu \geq 0$ *and* $\omega \leq \frac{L}{L-1}\nu$*, then there exists a network that is pairwise stable (except possibly for a set of agents with zero measure) where the proportion of agents of type* t *is equal to* π_t *for each* $t \in \mathcal{T}$ *if and only if there exists a vector of allocation parameters* α *satisfying Conditions 1 and 2 such that* $\pi_t = \frac{1}{\mu} \sum_H \mu_{v_1(t)} P_{H|v_1(t)} \alpha_H(t)$ *for every* $t \in \mathcal{T}$.

Proof. It suffices to show that specification (1) satisfies Assumption 3 when ν and ω are in the stated region. To establish Assumption 3(a), we compare the utility from adding a link to a distant individual of some type against the utilities from adding a link to individuals of the same type at each distance $\leq 2D$ (supposing this is feasible).

Without loss of generality, consider the possible changes in utility for an individual i of type t if a link is added to some individual k of type s. If $d(i,k;G) > 2D$, then type $\bar{t}$ is obtained. The change in utility is $f(X_i, X_k) + \epsilon_{il}(X_k) + |N(k)|\nu$, where $N(k)$ are k's existing friends (i.e., in type s). There is no intersection between $N(k)$ and i's existing friends of friends in type t, hence there are $|N(k)|$ additional friends of friends in type $\bar{t}$. If $d(i,k;G) = 2D = 4$, then some type in $\Psi(t,s)$ is obtained. This type yields the same utility as $\bar{t}$, because again there is no intersection between $N(k)$ and the existing friends of friends in type t (otherwise i and k would be at distance 3). If $d(i,k;G) = 3$, then k's friends and and i's friends of friends must intersect (this is how i and k are at distance 3). An example of this was shown in Figure 4. Hence there are fewer than $|N(k)|$ additional friends of friends in the resulting type, and so the change in utility is less than for type $\bar{t}$ (given $\nu \geq 0$).

Finally, suppose that $d(i,k;G) = 2$. In this case the change in utility from adding a link to k would include the value of mutual friendship (ω). Specifically, the marginal utility would include a term $\sum_{l=1,\, j(l)\neq\emptyset}^{L} G(j(l),k)\omega$ for the value of the new mutual friendships that are created by adding this link.[35] However for each new mutual friendship there would also be one fewer new friend of a friend, compared with the result of adding a link to a distant individual of type s. (Note that k's neighbors would be either a mutual friend or a friend of a friend to i in the resulting type, but not both.) In addition, individual k would no longer be a friend of friend to i. So there would be $\sum_{l=1,\, j(l)\neq\emptyset}^{L} G(j(l),k) + 1$ fewer friends of friends in the resulting type, call it $\hat{t}$, compared with type $\bar{t}$. Hence the difference between the utilities of these types is

$$u(\hat{t};\epsilon) - u(\bar{t};\epsilon) = \sum_{l=1, j(l)\neq\emptyset}^{L} G(j(l),k)(\omega - \nu) - \nu.$$

The maximum possible number of new mutual friendships is $L-1$ (because type t must have at least one link to spare), so having $\omega \leq \frac{L}{L-1}\nu$ ensures that this difference is weakly negative: $(L-1)(\omega-\nu) - \nu \leq 0$.

We have thus established that for each distance $d \leq 2D$, the utility of adding a link to someone of type s at that distance (if feasible) is weakly less than the utility of adding a link to someone of type s beyond $2D$. Therefore Assumption 3(a) is satisfied.

Now we show that Assumption 3(b) is satisfied in specification (1) in any region of the parameter space. The first line of (1) is always additively separable, whether or not the direct friends are in separate components of the graph without the ego. The second line is separable

[35] There could be multiple paths of length 2 from i to k, so there could be multiple new mutual friendships.

because there is no intersection among the friends of friends in separate components of the graph without the ego. To see how this delivers the required additive separability, let J and K contain the direct friends in two different components of the graph without the ego. Then

$$(\cup_{j\in J}N(j) - N(i) - \{i\}) \cap (\cup_{k\in K}N(k) - N(i) - \{i\}) = \emptyset,$$

and so

$$\left|\cup_{\hat{j}\in J\cup K}N(\hat{j}) - N(i) - \{i\}\right|\nu = |\cup_{j\in J}N(j) - N(i) - \{i\}|\,\nu + |\cup_{k\in K}N(k) - N(i) - \{i\}|\,\nu.$$

The third line of (1) is additively separable because $G(j,k) = 0$ for any pair of direct alters $j \in J$ and $k \in K$ (with J and K defined as above). Hence

$$\sum_{\hat{j}\in J\cup K}\sum_{\substack{\hat{j}'>\hat{j}\\ \hat{j}'\in J\cup K}} G(\hat{j},\hat{j}')\omega = \sum_{j\in J}\sum_{\substack{j'>j\\ j'\in J}} G(j,j')\omega + \sum_{k\in K}\sum_{\substack{k'>k\\ k'\in K}} G(k,k')\omega.$$

Therefore Assumption 3(b) is satisfied as well.

□

This result applies in both simulation exercises in Section 7, given the parameterizations that are used. In the first we have $\omega = \nu = 0$ and in the second we have $\omega = \nu = 0.2$. Hence in both cases we recover the sharp identified set.

C Statistical Inference

We take a sampling approach to inference whereby the target population is a single, large network (e.g., a school) and the statistical randomness (to be summarized through standard errors) comes from the I.I.D. sampling assumption (think here of choosing a set of students at random from the school). Naturally a different sampling procedure from the large network would lead to a different characterization of the statistical uncertainty, and here we illustrate our ideas with simple random sampling. See Kolaczyk (2009) for various sampling procedures in networks.

Proof of **Proposition (1)**: Let π be a scalar constant which represents the proportion of interest. (Here we show the proof for the scalar case for simplicity.) The sampling approach to inference maintains that we have a realization of the random process that generates a single, large network. Conditional on this population network, a simple random sample without replacement is drawn from the network.[36] Observations in this sample are by definition I.I.D. conditional on the network, and the only randomness in this second stage is coming from the act of sampling conditional on the realization of the large network from the random process. We take the population of interest to be a network of size N_p (unobserved) where we maintain that this population is large (but countable). Here we restrict ourselves to a finite population for simplicity, but similar results would hold for a more complex population. We subscript N with a p to indicate the population. Let there be a *sampled* network of size N_s from the population network where an observation in this sample is chosen using simple random sampling. This means that we choose a sample of size N_s from the population of size N_p conditional on the target (population) network. We enumerate the population by $\mathbf{T}'_p = (T_1, \ldots, T_{N_p})$ and the observed sample as $\mathbf{T}'_s = (T^s_1, \ldots, T^s_{N_s})$ where the sample is a subset from $\mathbf{T}_p$. We can take the $T's$ to be for example whether a given individual has any Black friends. The sampling process is conditional on the outcomes (the T's) and so the probability of choosing a sample of size N_s from N_p is $1/\binom{N_p}{N_s}$. Now define a sequence of binary indicator variables $(W_i)_{i=1}^{N_p}$. These are random variables such that $W_i = 1$ if observation i in the population is observed in the sample. We have: $E[W_i|\mathbf{T}_p] = P(W_i = 1|\mathbf{T}_p) = \frac{\binom{N_p-1}{N_s-1}}{\binom{N_p}{N_s}} = \frac{N_s}{N_p}$. Also, the W's are correlated with correlation $E[W_i W_j|\mathbf{T}_p] = \frac{N_s}{N_p}\frac{N_s-1}{N_p-1}$ for $i \neq j$ (the event that i is in the sample is not independent from whether $j \neq i$ is in the sample since the sample must always be of size N_s). Here, note that the expectation and the probability is conditional on the T's and so this randomness is purely from the simple random sampling process where $\mathbf{T}_p$ is taken as a vector of constants (i.e., the T's can be arbitrarily correlated). So, our estimator for π is:

$$\hat{\pi} = \frac{1}{N_s}\sum_{i=1}^{N_s} T^s_i = \frac{1}{N_p}\sum_{i=1}^{N_p} \frac{W_i}{\frac{N_s}{N_p}} T_i$$

First, it is simple to see that $\hat{\pi}$ is unbiased for π :

$$E[\hat{\pi}] = \frac{1}{N_p}\sum_{i=1}^{N_p} \frac{E[W_i|\mathbf{T}_p]}{\frac{N_s}{N_p}} T_i = \frac{1}{N_p}\sum_{i=1}^{N_p} T_i = \pi$$

[36]This is simpler but not essential as the difference between sampling with replacement is small especially in cases where the sample is much smaller than the population.

Now, for the variance, let σ_p^2 be the population variance: $\frac{1}{N_p}\sum_{i=1}^{N_p}(T_i-\pi)^2$.

$$\begin{aligned} Var(\hat{\pi}) &= \frac{\sigma_p^2}{N_s} - \frac{\sigma_p^2}{N_p} - \frac{1}{N_p^2}\sum_{i\neq j,i,j=1}^{N_p}(T_i-\pi)(T_j-\pi) \\ &\sim \frac{\sigma_p^2}{N_s} \end{aligned}$$

where the second line holds if indeed N_p is much larger than N_s and so the last two terms are negligible. Note here that the estimator is unbiased and its variance is proportional to $\frac{1}{N_s}$, the inverse of the sample size. So, as N_s increases, the variance approaches zero. In this finite population, N_s becomes closer to N_p, and the variance will be small. Note that the notion of "consistency" still holds in the finite population setup in that as N_s approaches N_p we learn the parameter of interest exactly since now (in a finite population) $\frac{1}{N_p}\sum_{i=1}^{N_p}T_i$ is the parameter of interest (rather than π) and here $P(W_i=1)$ approaches 1 as N_s approaches N_p (i.e., as everyone gets sampled). This is exactly Cochrane's definition of consistency in sampling whereby "estimates become exactly equal to the population value." See Cochrane (1977) p. 21.

An approximate $(1-\alpha)$ two sided confidence interval for π can then be constructed as $[\hat{\pi}-1.96\sqrt{\widehat{Var(\hat{\pi})}}, \hat{\pi}+1.96\sqrt{\widehat{Var(\hat{\pi})}}]$.[37] □

Asymptotic normality of sample means under simple random sampling from finite populations was studied under the approximate setup whereby both N_s and N_p approach infinity and results can be used to show that $N_s^{-\frac{1}{2}}\sum_{i=1}^{N_s}(T_i^s-\pi) \to^d \mathcal{N}(0,\pi(1-\pi))$. For example, Hajek (1960) provided necessary and sufficient conditions for the normal approximation to hold in sampling from finite populations when both N_s and N_p approach infinity. Essentially, he showed that a Lindberg-type condition must hold for the normal approximation to be valid. Here this condition holds trivially since the T's are binary. See Cochrane (1977) and Hajek (1960).

We now discuss how we can use the confidence intervals to map the uncertainty in sampling to θ which is the vector that characterizes the payoff structure. The data are informative only on the measure of network types, $\pi \equiv (\pi_t)_{t\in\mathcal{T}}$ – see Proposition 1 above.

Let there be a given vector π of observed type probabilities. Then, the identified set

[37] Again, the coverage for this interval holds only approximately for large N_p and N_s. In principle, one can use better approximations to the sampling distribution of the sample mean, but this is given here for simplicity.

$\Theta \subset \mathrm{R}^k$ in a given (large) network can be defined as follows (without conditioning on $\mathbf{X}$):

$$\Theta \equiv \Theta(\pi) = \{\theta \in R^k : F(\theta, \pi) = 0\}$$

where

$$F(\theta; \pi) = \min_{\{\alpha_{H(\theta)}(t) : t \in H(\theta)\}} \alpha^\top Q \alpha$$

subject to

$$\sum_{t \in H} \alpha_{H(\theta)}(t) = 1, \forall H(\theta)$$

$$\alpha_{H(\theta)}(t) \geq 0, \forall t, H(\theta)$$

$$\sum_{H} P_{H(\theta)} \alpha_{H(\theta)}(t) = \pi_t, \forall t$$

(see Section 6 for more on the quadratic matrix Q). Again, the key here is that if we know π, then constructing Θ becomes a family of quadratic programming problems, i.e., Θ collects all θ's where $F(\theta, \pi) = 0$. To obtain a confidence region for Θ, we can first construct a confidence region for π and then for every element π_s in this confidence region, solve for the corresponding Θ_s and take the union $\cup_s \Theta_s$. This heuristic relies fundamentally on being able to construct a valid confidence region for π. A sampling approach to inference delivers such a confidence region by sampling nodes independently within one large network conditional on outcomes. In this approach, the target parameter of interest, and the objective of the analysis would be to learn about the population type shares (and then using those to back out the structural parameters).

Take then as given that we have an approximation for the distribution of the vector of type shares. The previous discussion provides a sampling approach that can be used to obtain such an approximation.[38] Given the (approximate) distribution of types, we use standard methods to provide a confidence region for the identified set of the structural parameters.

A sample analog of the measure of each type is:

$$\hat{\pi}(t_k) = \frac{1}{n} \sum_i 1[i \in t_k]$$

[38] Other approaches may be possible. For example, the Bayesian bootstrap can be used to approximate via simple simulations the posterior for the vector of types using draws from gamma distributions.

for $k = 1, \ldots, |\mathcal{T}|$ and where these types are mutually exclusive. Moreover, let $\hat{\pi}^\top = (\hat{\pi}(t_1), \ldots, \hat{\pi}(t_{|\mathcal{T}|}))$ which is the vector of estimated type probabilities. We make the following assumption on the population choice probabilities and also maintain the approximate asymptotic distribution for the type vector. This result can easily be derived under standard assumptions (e.g., Theorem 17.2 in van der Vaart (1998)).

Assumption 4. *Let the network type proportions be such that*

$$\pi(t_k) > 0 \quad \forall k = 1, \ldots, |\mathcal{T}|; \quad \sum_{k=1}^{|\mathcal{T}|} \pi(t_k) = 1.$$

Also, assume that, as $n \to \infty$

$$G(\hat{\pi}, \pi) = n \sum_{k=1}^{|\mathcal{T}|} \frac{(\hat{\pi}(t_k) - \pi(t_k))^2}{\hat{\pi}(t_k)} \to_d \chi^2_{|\mathcal{T}|-1}, \tag{4}$$

where $\pi^\top \equiv (\pi(t_1), \ldots, \pi(t_{|\mathcal{T}|}))$.

Again, the exact asymptotic distribution as stated in (4) is one way to characterize sampling uncertainty and is not strictly required.

Given the above assumption, to build a (frequentist) confidence region for Θ, we "invert" the above statistic for multinomial probabilities. In particular, define

$$CI_{1-\alpha}(\pi) = \{\pi \in S^{|\mathcal{T}|} : G(\hat{\pi}, \pi) \leq c_{1-\alpha}(\chi^2_{|\mathcal{T}|-1})\} \tag{5}$$

where $S^{|\mathcal{T}|}$ is the unit simplex of size $|\mathcal{T}|$, $\hat{\pi}(t)$ are the sample analogues of the type probabilities, $c_{1-\alpha}(\chi^2_{|\mathcal{T}|})$ is the $(1-\alpha)$ critical value of the $\chi^2_{|\mathcal{T}|}$ distribution. The confidence region in (5) is standard and collects the set of network type probabilities that covers the truth with probability $(1 - \alpha)$ (in repeated samples). It is also possible to consider a Bayesian approach to inference here where obtaining a posterior for $\pi(t)$ given standard priors can be easily done also (using a Bayesian bootstrap, for example).

Now, for every $\pi \in CI_{1-\alpha}(\pi)$, we can solve our model in terms of the *set* of θ's using the quadratic programming function $F(\theta, \pi) = 0$. The collection of these sets would be a confidence region for the identified set:

$$CI_{1-\alpha}(\theta) = \{\Theta(\pi) : F(\Theta(\pi), \pi) = 0 \quad \text{for} \quad \pi \in CI_{1-\alpha}(\pi)\} \tag{6}$$

(row)	(parameter)																
1	$\alpha_1(B,0)$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	$\alpha_2(B,0)$	0	**1**	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0
3	$\alpha_2(B,B)$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	$\alpha_3(B,0)$	0	0	0	0	0	0	0	0	0	0	0	**1**	0	**1**	0	0
5	$\alpha_3(B,W)$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	$\alpha_4(B,0)$	0	**1**	0	0	0	**1**	0	0	0	0	0	**1**	0	**1**	0	0
7	$\alpha_4(B,B)$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	$\alpha_4(B,W)$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	$\alpha_5(W,0)$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	$\alpha_6(W,0)$	0	0	0	0	0	0	0	0	0	**1**	0	0	0	**1**	0	0
11	$\alpha_6(W,W)$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	$\alpha_7(W,0)$	0	0	0	**1**	0	**1**	0	0	0	0	0	0	0	0	0	0
13	$\alpha_7(W,B)$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	$\alpha_8(W,0)$	0	0	0	**1**	0	**1**	0	0	0	**1**	0	0	0	**1**	0	0
15	$\alpha_8(W,W)$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	$\alpha_8(W,B)$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 5: Matrix Q for Model with $D = 1$, $L = 1$, and $\mathcal{X} = \{B, W\}$

Here, the notation for $\Theta(\pi)$ in $F(\Theta(\pi), \pi) = 0$ implicitly means that $\Theta(\pi)$ is the set of θ's such that $F(\theta, \pi) = 0$. Then we can easily show the following theorem.

Theorem 4. *Let Assumption 4 above hold. Then, we have*

$$\lim_{n \to \infty} Pr\{\Theta \subset CI_{1-\alpha}(\theta)\} = 1 - \alpha. \tag{7}$$

This result is obtained from noticing that $\Theta(\pi) \subset CI_{1-\alpha}(\theta) \Leftrightarrow \pi \in CI_{1-\alpha}(\pi)$. Hence, $Pr\{\Theta \subset CI_{1-\alpha}(\theta)\} = Pr\{\pi \in CI_{1-\alpha}(\pi)\}$, which converges to $1 - \alpha$. For other promising approaches to inference, see for example the results in Leung (2015).

D Details of Simulation Procedures

D.1 Objective Matrix in First Exercise

The matrix Q for the first simulation exercise is shown here in Appendix Figure 5. The rows and columns correspond to the allocation parameters listed under the heading "parameter."

Table A1: Equilibrium Type Shares in the First Simulation

Type $v = (x, y)$	Proportion (conditional on race of the ego)
$(B, 0)$	$\pi_{(B,0)} = (1 - f_{BB})(1 - f_{BW}) + (1 - \alpha_3(B, W))(1 - f_{BB})f_{BW}$
(B, B)	$\pi_{(B,B)} = f_{BB}(1 - f_{BW}) + (1 - \alpha_4(B, W))f_{BB}f_{BW}$
(B, W)	$\pi_{(B,W)} = \alpha_3(B, W)(1 - f_{BB})f_{BW} + \alpha_4(B, W)f_{BB}f_{BW}$
$(W, 0)$	$\pi_{(W,0)} = (1 - f_{WW})(1 - f_{WB}) + (1 - \alpha_7(W, B))(1 - f_{WW})f_{WB}$
(W, W)	$\pi_{(W,B)} = f_{WW}(1 - f_{WB}) + (1 - \alpha_8(W, B))f_{WW}f_{WB}$
(W, B)	$\pi_{(W,W)} = \alpha_7(W, B)(1 - f_{WW})f_{WB} + \alpha_8(W, B)f_{WW}f_{WB}$

D.2 Simplification of the QP Problem in First Exercise

As noted in the text, the QP problem for this model can be simplified to the point that it is trivial to verify whether the optimal value is zero. We use this result to confirm the identified set obtained using the MCMC search procedure with the original QP problem, and to construct identified sets based on different observations (i.e., different type shares).

The simplification is obtained as follows. There are 16 potentially nonzero allocation parameters in the QP problem for this model (listed in Appendix Figure 5), but 12 can be eliminated with simple manipulations. (Specifically, the four allocation parameters with a positive diagonal element in their row of the matrix Q are set equal to zero, as this is necessary for an optimal value of zero to be attainable, and eight other parameters are eliminated using the constraint $\sum_{t \in H} \alpha_H(t) = 1$.) Expressions for the equilibrium type shares as a function of the remaining four allocation parameters and the structural parameters are listed in Table A1. Unique values of these remaining allocation parameters can then be recovered, given a vector of structural parameters and the vector of type shares. It is then trivial to compute the objective function value and to assess whether these allocation parameters satisfy the constraint $0 \leq \alpha_H(t) \leq 1$.

It is also possible to use the expressions in Table A1 to find all the equilibrium type shares for a given vector of structural parameters. Rather than evaluate different parameter vectors given a fixed vector of type shares (to find the identified set), one can instead evaluate different type shares given a fixed vector of structural parameters. Either way, for any pair of vectors of type shares and structural parameters, one recovers the four allocation parameters using the expressions in the table and verifies whether the objective function value is zero and the allocation parameters each fall within the unit interval. Then, because Conditions 1 and 2 are necessary *and* sufficient for pairwise stability in this model (because $D = 1$, see Appendix B), this guarantees that the type shares are obtainable in equilibrium under the given values of the structural parameters. (There is one additional restriction on the admissible vectors of type shares, which is that the measure of blacks linked to whites must equal the measure of whites linked to blacks.) Thus, to find the set of equilibria shown in Figure 6, we fix the structural parameters at the stated values and evaluate a grid of points in the space of admissible vectors of type shares.

D.3 Utility Specificaton in the Second Exercise

This exercise uses $D = 2$, $L = 3$, and $\mathcal{X} = \{B, W\}$. Written in terms of the matrix-vector pairs (A, v) that represent network types, the utility specification (1) is as follows:

$$\begin{aligned} u(A, v; \epsilon) \equiv & \sum_{l=2}^{L+1} a_{1l}\big(f_{v_1,v_l} + \epsilon_{l-1}(v_l)\big) && \text{(direct connections)} \\ & + \nu \sum_{k>L+1} 1\Big\{\sum_{l=2}^{L+1} a_{1l} a_{lk} > 0\Big\} && \text{(friends of friends)} \\ & + \omega \sum_{l=2}^{L+1} \sum_{k>l}^{L+1} a_{1l} a_{lk} && \text{(mutual friends)} \end{aligned} \tag{8}$$

(recall that row 1 of A corresponds to the ego, rows 2 to $L+1$ correspond to direct connections, and rows $k > L + 1$ correspond to friends of friends).

The order of the direct alters in the first line above is somewhat arbitrary because it is set by the convention we adopt to select a canonical representation for each network type (see Appendix D.5.1). This raises a question of how to assign the shocks for the direct connections. Rather than associate each shock with the same row of A and element of v in every type, we instead use the highest valued shocks (within each race) to compute the utility of each type. Then, for example, the utility of a type with one black friend and one white friend does not depend on whether the black alter corresponds to row 2 of A and the white alter corresponds to row 3, or vice versa. More generally, by assigning the shocks in this way, the utility of each network type does depend on the particular convention used to select the canonical representations.

D.4 Microsimulation Procedure

To simulate vectors of type shares to use as data, we generate equilibrium networks from which the shares can be extracted. The description below of this microsimulation procedure is focused on the second exercise, but the overall procedure is the same for both.

Although only one network is needed for the type shares, we generate a number of pairwise stable networks to illustrate the variation that can arise in these models. In the second exercise, each network has $n = 500$ individuals, with $n_B = 100$ blacks and $n_W = 400$ whites. For each network we first draw vectors of preference shocks for all the individuals. Then the procedure to find a pairwise stable network starts with a random initial graph. These initial graphs are generated by independently establishing links with probability $1/(2n)$ and then removing links at random from individuals with more than L links. The success rate of $1/(2n)$ is chosen to limit the number of individuals with greater than L links in the initial draw while yielding a degree distribution that is somewhat similar to the equilibrium distribution.

Given a random initial graph, the following sequential process is then used to find a stable network:

1. Draw a random sequence over all unordered pairs of players (i.e., a permutation of the numbers 1 to $n(n-1)/2$, which index the pairs);

2. For each pair (i, j) in the sequence, myopically update $g_{ij} = g_{ji}$ based on the conditions for pairwise stability, using the network as it has evolved up to that point;

3. If no links or non-links were updated in an entire sequence over all the pairs, stop: the network is pairwise stable;

4. Otherwise go through another random sequence of all pairs: repeat steps 1 to 3, up to `#seqs` times (`#seqs` was set to 100).

If the network does not converge after `#seqs` of such random sequences over all pairs, a new random initial graph is used and steps 1 to 4 are repeated. If the network still does not converge after this process is repeated with multiple initial graphs (up to seven), we say we have failed to find an equilibrium for this set of preference shocks and move on to draw a new set of shocks. Networks that do not converge are discarded.

We generated a total of 47 pairwise stable networks in this way (out of 50 attempts). The degree distribution from these networks appears in Figure A5, along with the degree distribution of same-sex friendships from all schools in the Add Health data. Our simulated networks have fewer isolates and more individuals with one link, but otherwise the two distributions are broadly similar, and the average degree is the same at 1.05. Also, it turns out that the shares of network types with any mutual friends are zero in most simulated networks (see Figure 9). This is a consequence of having $\omega = \nu$ along with the values of the other parameters that were chosen to generate a degree distribution like that in Add Health. Under these parameter values there is very low probability that three randomly selected individuals would all desire to be connected with each other in a triad, rather than at least one of them preferring to drop one link (thereby gaining ν while losing ω and $f_{xy} + \epsilon_{l-1}(y)$).[39]

D.5 Network Types and Transformations of Types

Our approach requires an enumeration of all possible network types under a given preference structure. In addition we define two sets of transformations on the types: what they become if one of their links is deleted, and what they become if a link is added to some other type. These transformations are needed to generate preference classes and to assess Conditions 1 and 2. None of these depend on a particular parameterization of the model, so they can be constructed prior to the recovery of the identified set. Hence any computational burden here does not directly impact the time it takes to search through the parameter space.

D.5.1 Enumeration of Network Types

Each network type is an equivalence class of isomorphic subnetworks (with a root node, which is the ego). There can be multiple matrix-vector pairs (A, v) representing the same type, which are related to each other by permutation of the rows and columns for the alters.[40] So for computational convenience, we adopt a convention to single out one (A, v) pair from each class, which we refer to as the canonical representation of that network type. The enumeration of network types is then a list of these canonical representations.

Our convention is as follows. After the first line (row/column of A or element of v),

[39] Given $\omega = \nu$, the marginal payoff from dropping one link in a triad of players (with no other connections) is $-(f_{xy} + \epsilon_l(y))$. The highest value of any f_{xy} is $f_{WW} = -0.7$, and with this value the probability that a white individual would *not* want to drop such a link to the alter at position l is $\Pr(f_{WW} + \epsilon_l(W) > 0) = 0.24$. Among three arbitrary whites, the probability that *none* of them would want to drop either one of their two links in the triad is $\Pr(f_{WW} + \epsilon_l(W) > 0)^6 = 0.0002$. This, therefore, is the probability that a triad (with no other connections) would be pairwise stable among three randomly selected whites.

[40] In general, finding whether or not two graphs are isomorphic has an unknown computational complexity. It is known to be in **NP**, but not whether it is in **NP**-complete or **P**.

which corresponds to the individual of interest (the ego), the next L lines correspond to her direct connections (or direct alters). Then the subsequent $L-1$ lines correspond to the $L-1$ additional possible direct connections of the first direct alter, and so on. Should the ego have fewer than L links, we leave vacant lines at the end of her block. For example, if the ego only has $L-1$ links, the $L+1$ row and column of A and element of v are zero. This also applies to the blocks for any alter who does not have his full set of links. Second, if an indirect alter is reached through multiple direct alters, she appears in the block corresponding to the direct alter with the most links. Finally, an ordering over the set of characteristics $\mathcal{X}$ in the vector v fixes the permutation and selects the canonical element from the equivalence class of (A, v) pairs for this type.

Then, given such a convention, it is useful to have an automated procedure to generate the list of canonical representations. First we generate all non-isomorphic adjacency matrices A. This is similar to generating all unlabeled graphs with up to $1 + L\sum_{d=1}^{D}(L-1)^{d-1}$ nodes (10 nodes in the model with $D = 2$ and $L = 3$). Various algorithms for graph generation are available.[41] However for this model, given the limitation on the number of links per node ($L = 3$), it was easiest to write our own simple procedure to generate the non-isomorphic adjacency matrices. First we make all the tree structures (i.e., graphs with no cycles), then all graphs with one mutual friendship, then all with two mutual friendships, and finally the one graph with three mutual friendships. There are a total of 36 non-isomorphic adjacency matrices that are relevant under this preference structure.[42]

Finally to construct the network types we consider all possible combinations of characteristics of the ego and the direct alters. (The characteristics of the alters at distance 2 are not relevant under this preference structure, so they can be omitted from the vectors v.) We then compare permutations of the alter characteristics and retain only those (A, v) pairs that are unique, following our convention.[43] This yields the list of canonical representations of network types. In the model for the second exercise there are 356 distinct network types.

[41] See, for example, `http://www3.cs.stonybrook.edu/~algorith/files/generating-graphs.shtml` for a list with recommendations.

[42] This is considerably less than the number of unlabeled graphs among 10 nodes for three reasons. First, here the nodes have at most three links. Second, we restrict to graphs with one connected component (which contains the ego). Third, we do not consider links among nodes at distance 2 from the ego, as they are not relevant for the ego's utility.

[43] General algorithms to test for isomorphisms between graphs with node characteristics (i.e., "colors") are also available. For example, the nauty and Traces programs (`http://pallini.di.uniroma1.it/`).

D.5.2 Link Deletion

The construction of preference classes involves comparing the utility of each type against what would be obtained if a link were deleted. To facilitate these comparisons, we make a list containing the results of link deletion from each type. Links are easily deleted from a network type by setting the relevant elements of A to zero. We do not then need to check which canonical representation is isomorphic to the result; only the utility of the resulting type is needed. Utility in specification (1) is computed as a function of the characteristics of the direct alters, the number of friends of friends, and the number of mutual friends. These are easily extracted from any (A, v) pair regardless of the ordering of the rows and columns, and nodes that are not connected to the ego following the deletion of a link can be ignored automatically. Accordingly, the list regarding link deletion contains only the characteristics of the direct alters and the numbers of indirect and mutual friends in the resulting types. The number of elements in the list is the number of types multiplied by the number of links. In the second exercise this is $356 \times 3 = 1068$.

D.5.3 Link Addition

In order to construct the objective matrix Q in our QP problem, we need a mapping that gives the types which would result if individuals of two types, say t and s, were linked (given an initial distance $> 2D$). This mapping can be stored as a matrix where each row and each column corresponds to a type, and the entry at position $[t, s]$ gives the type $\bar{t}$ that would result for an individual of type t if a link were added to an individual of type s. (The matrix is $|\mathcal{T}| \times |\mathcal{T}|$.) For any cases where either t or s already has L links, this entry is blank, which makes the matrix sparse. Otherwise, the resulting type $\bar{t}$ is found by: (1) adding a link in the first unoccupied row and column for a direct alter in the adjacency matrix for type t, A_t; (2) inserting the characteristic of the ego from type s into the corresponding element of the vector of characteristics for type t, v_t; and (3) adding links to indicate any direct alters in type s into the appropriate unoccupied rows for *indirect* alters in A_t. This yields an adjacency matrix and vector of characteristics representing the new type $\bar{t}$. The resulting (A, v) pair may not be the canonical representation of that type, however, so we apply an algorithm to test for graph isomorphisms to find the matching element within the list of canonical representations. We wrote our own simple algorithm, which considers certain permutations of A and v, but more general algorithms could be used.[44]

D.6 Specification of the QP Problem

In order to accommodate error in the type shares estimated from a finite sample, we modify QP problem (3) to allow the predicted type shares to be within fixed bands around the observed shares. To do this we define two slack variables for each type share, one for a positive difference, $\beta^+(t)$, and one for a negative difference, $\beta^-(t)$. The constraints for matching predicted shares to observed shares then become

$$\frac{1}{\mu}\sum_H \mu_{v_1(t)} P_{H|v_1(t)}(\theta)\, \alpha_H(t) + \beta^+(t) - \beta^-(t) = \pi_t,\ \forall t.$$

The slacks above are additional variables in the modified QP problem (although the objective function is unchanged). Their magnitudes are limited based on functions of the sample size n, denoted $\delta^+(n)$ and $\delta^-(n)$, by the following additional constraints:

$$0 \leq \beta^+(t) \leq \delta^+(n) \ \text{ and } \ 0 \leq \beta^-(t) \leq \delta^-(n),\ \forall t.$$

Thus the slacks define fixed bands around the observed type shares, allowing errors from $-\delta^-(n)$ to $+\delta^+(n)$ for each type.

The slacks are not minimized in the modified QP problem (e.g., by including their sum of squares in the objective function) because we found that doing so would add greatly to the solution time. However the sums of the slacks ($\sum_t \beta^+(t)$ and $\sum_t \beta^-(t)$) are further constrained with an upper bound. Without this, the total absolute error between the predicted and observed type shares could be equal to $\delta^+(n)$ or $\delta^-(n)$ multiplied by the number of types. Given the number of types in the second exercise, even small amounts for these errors like 0.01 could then result in a large total absolute error—greater than 1, for example, which would be the total absolute error if we just predicted each type share to be equal to zero.

[44] See footnote 43 for references. We only need to consider a limited number of permutations of A and v because the canonical representations always place the ego in the first row, the direct alters in the next L rows, and the indirect alters in specific rows based on the direct alter through which they are reached. Only permutations among blocks for direct alters and for indirect alters need to be checked.

Hence we include the constraints

$$\sum_t \beta^+(t) \le \Delta^+ \quad \text{and} \quad \sum_t \beta^-(t) \le \Delta^-.$$

The Δ^+ and Δ^- above, along with the $\delta^+(n)$ and $\delta^-(n)$, can be thought of as tuning parameters. The specific values we choose are described in Appendix D.7.3.

The exact formulation of the QP problem used in our simulations is then as follows:

$$\min_{\{\alpha_H(t): t \in H, H \in \mathcal{H}(\theta)\}, \beta^+, \beta^-} \alpha^\top Q(\theta)\, \alpha, \quad \text{subject to:}$$

$$\frac{1}{\mu} \sum_{H \in \mathcal{H}(\theta)} \mu_{v_1(t)} P_{H|v_1(t)}(\theta)\, \alpha_H(t) + \beta^+(t) - \beta^-(t) = \pi_t, \ \forall t \tag{9}$$

$$\sum_{t \in H} \alpha_H(t) = 1, \ \forall H \in \mathcal{H}(\theta); \qquad 0 \le \alpha_H(t) \le 1 \tag{10}$$

$$0 \le \beta^+(t) \le \delta^+(n), \ \forall t \tag{11}$$

$$0 \le \beta^-(t) \le \delta^-(n), \ \forall t \tag{12}$$

$$\sum_t \beta^+(t) \le \Delta^+ \tag{13}$$

$$\sum_t \beta^-(t) \le \Delta^- \tag{14}$$

The dependence of the objective matrix and the set of preference classes on the structural parameter vector (i.e., $Q(\theta)$ and $\mathcal{H}(\theta)$), is a further aspect of our implementation, discussed in the next section.

D.7 Evaluation of a Parameter Vector

Given a candidate vector of preference parameters θ, we wish to solve the QP problem above to determine whether the optimal value is zero. There are three main steps in this process: 1) finding the distribution of preference classes, 2) constructing the objective matrix for the QP problem, and 3) solving the QP problem.

D.7.1 Distribution of Preference Classes

The probability distribution of preference classes is approximated by Monte Carlo integration with independent draws of the preference shock vectors (we used 10000 draws). For each draw ϵ_i we find the preference class of a black individual and a white individual with those particular shocks. These are the two sets of types such that $u(A, v; \epsilon_i) \geq u(A_{-l}, v; \epsilon_i)$, $1 \leq l \leq L$, given $v_1 = B$ (a black ego) and given $v_1 = W$ (a white ego). The number of times a particular preference class appears with different draws then approximates its true probability. Thus we have the preference class probabilities, $P_{H|x}(\theta)$, and the set of preference classes that have appeared in this procedure, $\mathcal{H}(\theta)$. As the notation indicates, the contents of $\mathcal{H}(\theta)$ can change with θ, because many preference classes have low probabilities of occurrence and so may not be realized even with $10,000$ draws, depending on the values of θ.

To give a sense of the magnitudes of these sets, at the true parameter values we generate 249 preference classes (i.e., $|\mathcal{H}(\theta_\circ)| = 249$). The number of potentially nonzero allocation parameters (which are variables in the QP problem) is equal to the sum of the cardinalities of these preference classes: $\sum_{H \in \mathcal{H}(\theta_\circ)} |H| = 5,013$. These are the results when we restrict to network types that are either observed in the data or adjacent via addition or deletion of a link (all other types can be ignored). If we do not remove the unobserved and non-adjacent types (216 out of the 356 types in this model), we would generate 278 preference classes with a total of 12,812 potentially nonzero allocation parameters.

D.7.2 Construction of the QP Objective Matrix

Section 6 gives an overview of the construction of the objective matrix Q. Here we provide some additional detail on the construction of the matrix S (the precursor to Q). Each row in S corresponds to an allocation parameter, as does each column. The row for parameter $\alpha_H(t)$ indicates which allocation parameters (in the columns) correspond to individuals of types that someone of type t with preferences in class H would like to add a link to. More specifically, the entries of S are defined as $S_{[\alpha_H(t), \alpha_G(s)]} = 1_{\bar{t}(s) \in H}$, where type $\bar{t}(s)$ is the type that an individual of type t (from $\alpha_H(t)$, for the row) would become if they added a link to someone of type s (from $\alpha_G(s)$, for the column) at a distance greater than $2D$.

To construct S, we first extract and store the sequence of types associated with all the allocation parameters in the columns (e.g., the type s in $\alpha_G(s)$). In practical terms, this is the concatenation of the contents of all the preference classes in $\mathcal{H}(\theta)$. We then proceed by row, as follows. Given the allocation parameter for the row, $\alpha_H(t)$, we firsts use the matrix defined in D.5.3 to find the type(s) s that someone of type t with preferences in class H would like to add a link to. As described in D.5.3, entry $[t, s]$ in that matrix indicates the type $\bar{t}$ that an individual of type t would become after adding a link to an individual of type s (or the entry is blank if either types t or s already have L links). Accordingly, we take row t of the matrix from D.5.3 and identify any entry whose value $\bar{t}$ is contained in H. The column positions of these entries then indicate the types s that someone of type t with preferences in class H would like to add a link to. Thus we have a list of desired alter types for these individuals of type t. This list is then compared with the sequence of types from the allocation parameters in the columns (e.g., the s in $\alpha_G(s)$). In columns where there is a match, the entries of this row for allocation parameter $\alpha_H(t)$ are set to 1. Those columns correspond to allocation parameters $\alpha_G(s)$ such that $\bar{t}(s) \in H$. The other entries in this row are set to 0. This yields the desired result: $S_{[\alpha_H(t), \alpha_G(s)]} = 1_{\bar{t}(s) \in H}$. The main advantage of this approach is that the match procedure can be applied to the entire row at once, and it runs quickly even though the sequence of types from the columns of S is large.

To save memory, S is stored as a sparse binary matrix. Also, because the contents of $\mathcal{H}(\theta)$ can change with θ (see D.7.1), the matrices S and Q are re-constructed for each candidate parameter vector θ. While this adds a small amount of computational time (relative to the time to solve the QP), it turns out to be much better for memory usage compared with trying to maintain a fixed list of preference classes and a constant version of the matrices. As noted earlier, many preference classes have very low probabilities and do not appear in the list $\mathcal{H}(\theta)$ that is generated from a particular vector θ. A fixed matrix Q that could accommodate all preference classes found with any vector in the parameter space would be vastly larger than the matrices $Q(\theta)$ that are constructed for particular values of θ.

D.7.3 Solution of the QP Problem

To speed the solution of the QP problem we first use a linear programming (LP) problem to obtain starting values for the allocation parameters, which are those that minimize the sum of absolute deviations between the observed and predicted type shares. The LP problem for this is similar to the QP problem presented in D.6, except for the objective function and an absence of upper bounds on the slack variables. It is specified as follows:

$$\min_{\{\alpha_H(t): t\in H, H\in\mathcal{H}(\theta)\},\beta^+,\beta^-} \sum_t \left(\beta^+(t) + \beta^-(t)\right)$$

subject to:

$$\frac{1}{\mu}\sum_{H\in\mathcal{H}(\theta)} \mu_{v_1(t)} P_{H|v_1(t)}(\theta)\, \alpha_H(t) + \beta^+(t) - \beta^-(t) = \pi_t,\ \forall t$$

$$\sum_{t\in H} \alpha_H(t) = 1,\ \forall H \in \mathcal{H}(\theta); \qquad 0 \le \alpha_H(t) \le 1$$

$$0 \le \beta^+(t), \beta^-(t),\ \forall t$$

The solution to this problem provides a vector of allocation parameters and slack variables that are used as starting values in the QP problem. Also the sums of the slacks in the solution here are used to define the limits Δ^+ and Δ^- in constraints (13) and (14) of the QP problem. Specifically we set the values of Δ^+ and Δ^- equal to $\max\left\{\frac{1}{2}\sum_t [b^+(t) + b^-(t)], 6/n\right\}$, where $b^+(t)$ and $b^-(t)$ are the slacks in the solution to the LP problem. This limits the sum of absolute errors in the QP problem to the (optimal) sum of absolute errors from the LP problem, but with a floor of $6/n$. The floor is required in order to maintain some minimal size for the bands around the observed type shares. Last, for constraints (11) and (12) of the QP problem, we set $\delta^+(n) = 2/n$ and $\delta^-(n) = 1/(2n)$. These are roughly the amounts required in order for the solver to converge easily when we use the true parameter values and the observed type shares from the one randomly selected network. These values and the floor of $6/n$ in the formula for Δ^+ and Δ^- function as tuning parameters, which can be adjusted based on the performance of the solver.

To solve the QP problem, we use the active set algorithm in the program KNITRO, which is a variant of a sequential linear and quadratic programming optimization method (Byrd, Gould, Nocedal, and Waltz 2003). As detailed below, this routine performs well on our problem. Over a range of values for the preference parameters, which yield on the order of 2,000 to 10,000 allocation parameters, the solution time averages less than 25 seconds.

D.8 Construction of the Identified Set

In concept, the identified set is a level set in the space of structural parameters, where the optimal value of the QP is zero and the predicted type shares match the observed type shares. Our approach to find this level set involves the use of Markov Chain Monte Carlo (MCMC) procedures. The results from the solution of the QP problem for a given parameter vector are converted into a pseudo-density, which an MCMC algorithm can then use to draw parameter vectors and move through the parameter space.[45]

We use a log pseudo-density that is proportional to $-[\alpha^*(\theta)^\top Q(\theta)\alpha^*(\theta) + \beta^*(\theta)^\top \beta^*(\theta)]$, where $(\alpha^*(\theta), \beta^*(\theta))$ denotes a solution to the QP problem for θ. This allows for small positive values of the objective function as well as the errors between observed and predicted type shares. We found this to be helpful to address issues in computational precision, along with the sampling error in the observed type shares. Structural parameter vectors where the value of this pseudo-density is at least 95% of its maximum are then considered to be in the identified set and hence are shown in the figures.[46] For the results plotted in Figures 10 and 11, we generated a total of 7,090 such vectors.

D.9 Computational Performance

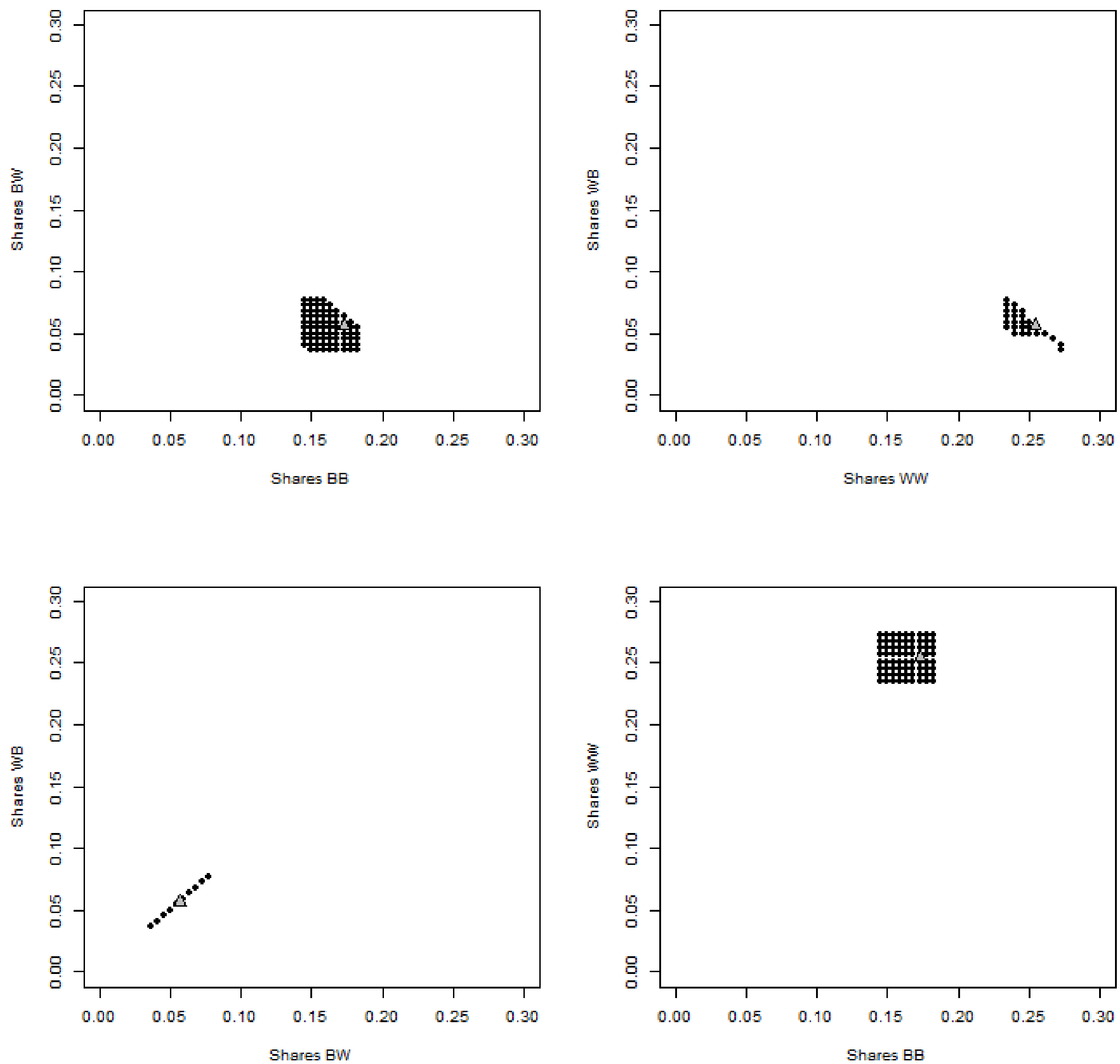
The above procedures to construct the identified set were run on machines with Intel® Xeon® 5160 processors (3 GHz base frequency) and 16 GB of physical memory. Computations were not parallelized, except in the "embarrassingly" simple sense that multiple Markov Chains were run on different machines. The time required to evaluate a single parameter vector θ consists mainly of three steps: generating the preference class distribution, constructing the objective matrix, and solving the QP problem (relative to these, the time to solve the preliminary LP problem is trivial). On average the first two steps each account for only 10% of the total compute time, so the majority of the computational burden comes from the solution of the QP (i.e., 80% of the compute time).

Based on evaluations of 15,000 structural parameter vectors in total, the average time to evaluate a single parameter vector (i.e., to generate the pseudo-density for a given θ) was 29.8 seconds. The number of allocation parameters in the QP problems for these parameter vectors ranged roughly from 2,000 to 10,000, with an average of 5,955.3.

[45] For the results shown here, we used both the Metropolis-Hastings and slice sampler algorithms in Matlab.

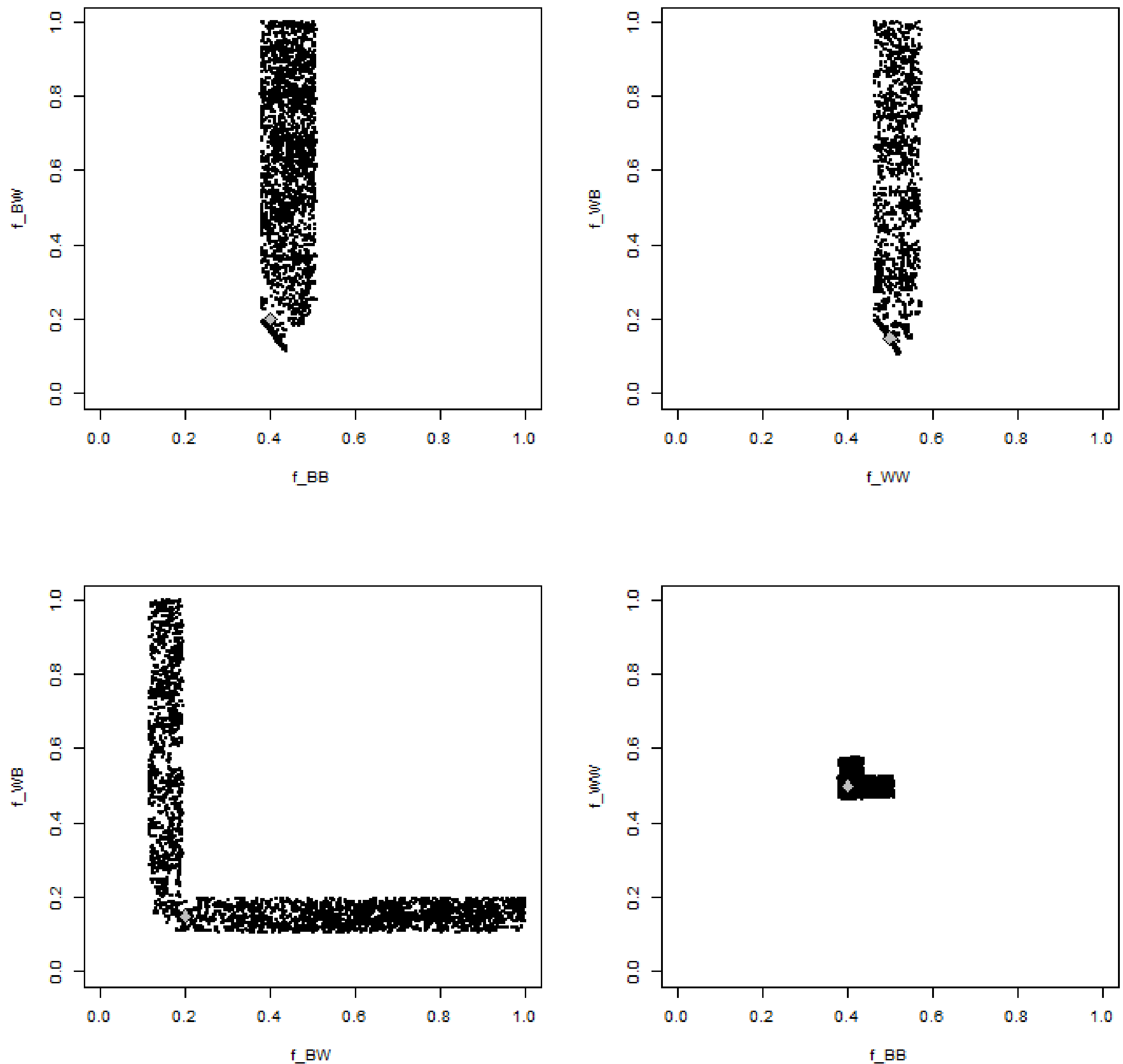
[46] The 95% threshold is somewhat arbitrary, but the results are robust to the value that is used. Relatively few parameter vectors produce pseudo-densities that are between 90% and 99% of the maximum value, for example.

Figure 6: Equilibrium Type Shares in First Simulation Exercise



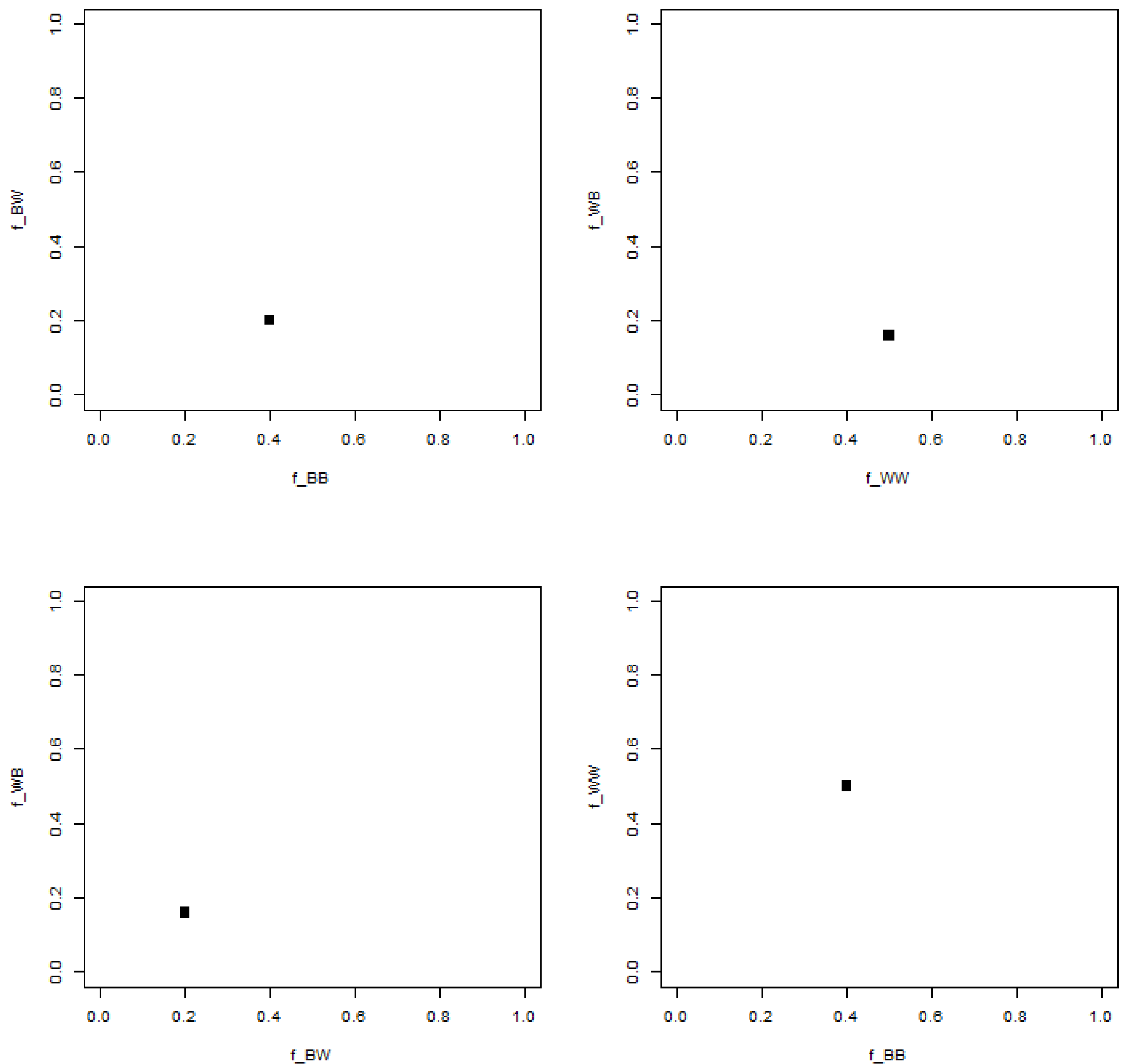
Notes: Points illustrate the full set of equilibrium type shares under parameter values $f_{BB} = 0.40$, $f_{BW} = 0.20$, $f_{WB} = 0.15$, and $f_{WW} = 0.50$. Triangles indicate the type shares from one finite network simulation, with values $\pi_{(B,B)} = 0.173$, $\pi_{(B,W)} = 0.056$, $\pi_{(W,B)} = 0.056$ and $\pi_{(W,W)} = 0.255$.

Figure 7: Identified Set from One Finite Network in First Exercise



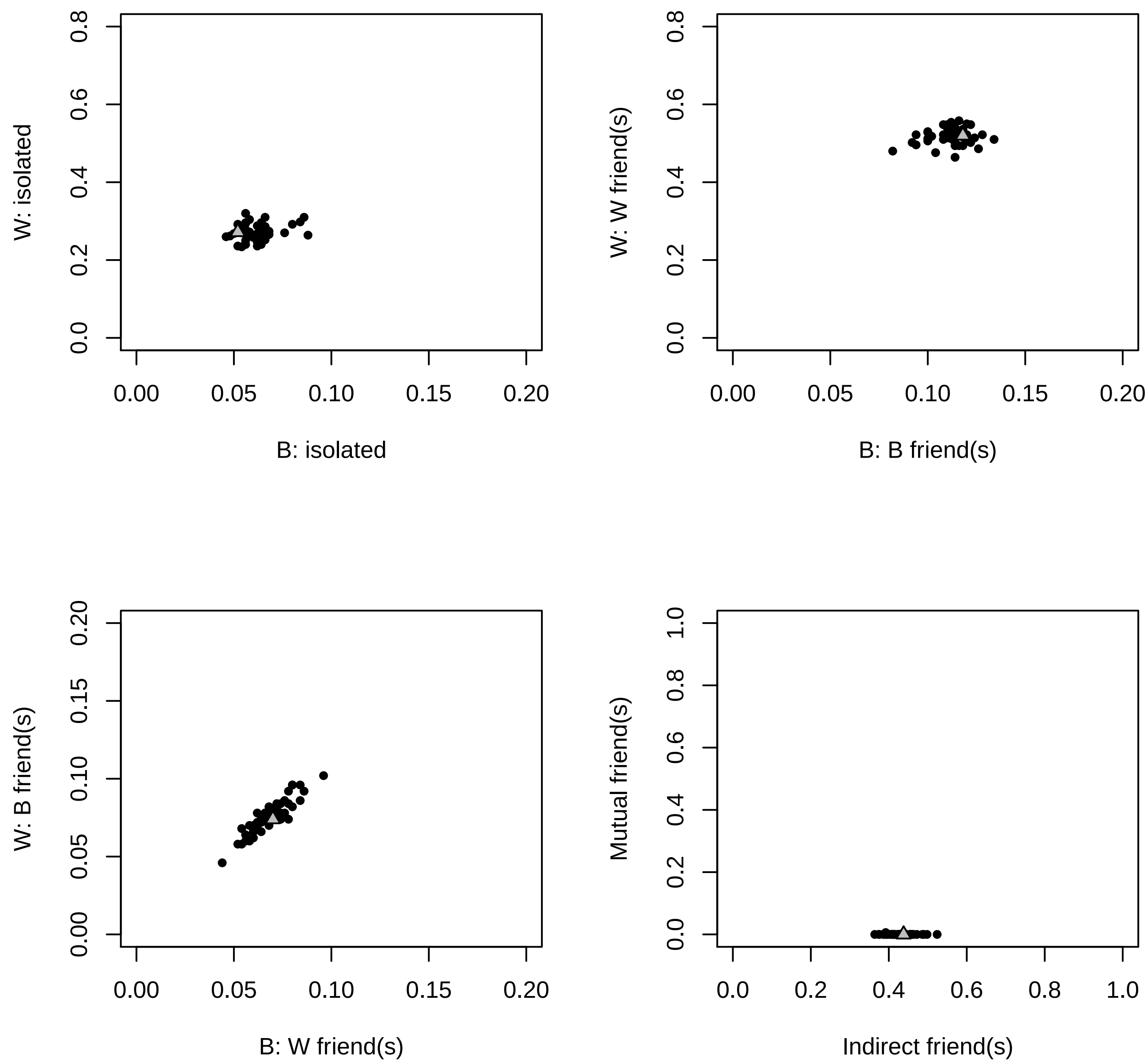
Notes: Points illustrate the identified set obtained using type shares from one finite network (shown with triangles in Figure 6). Diamonds indicate true parameter values: $f_{BB} = 0.40$, $f_{BW} = 0.20$, $f_{WB} = 0.15$, and $f_{WW} = 0.50$.

Figure 8: Identified Set from Four Randomly Selected Networks in First Exercise



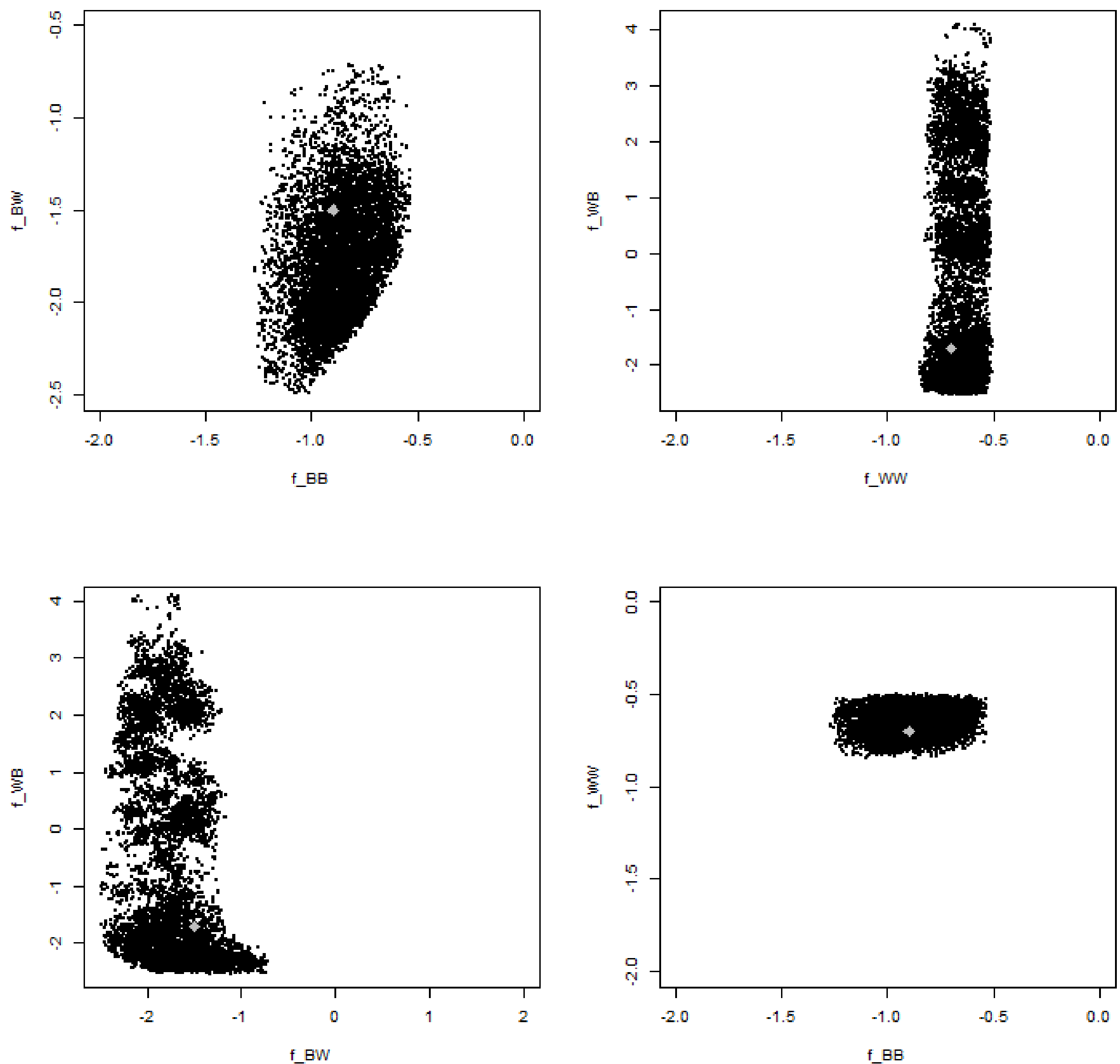
Notes: Blocks illustrate the identified set obtained using four vectors of type shares (shown in Figure A2) that were randomly selected from the full set of equilibrium type shares in Figure 6. A grid of parameter vectors with intervals of size 0.02 in each dimension was evaluated, and the identified set consists of one vector in this grid: $\hat{f}_{BB} = 0.40$, $\hat{f}_{BW} = 0.20$, $\hat{f}_{WB} = 0.16$, and $\hat{f}_{WW} = 0.50$. True parameter values are: $f_{BB} = 0.40$, $f_{BW} = 0.20$, $f_{WB} = 0.15$, and $f_{WW} = 0.50$.

Figure 9: Equilibrium Type Shares in Second Simulation Exercise



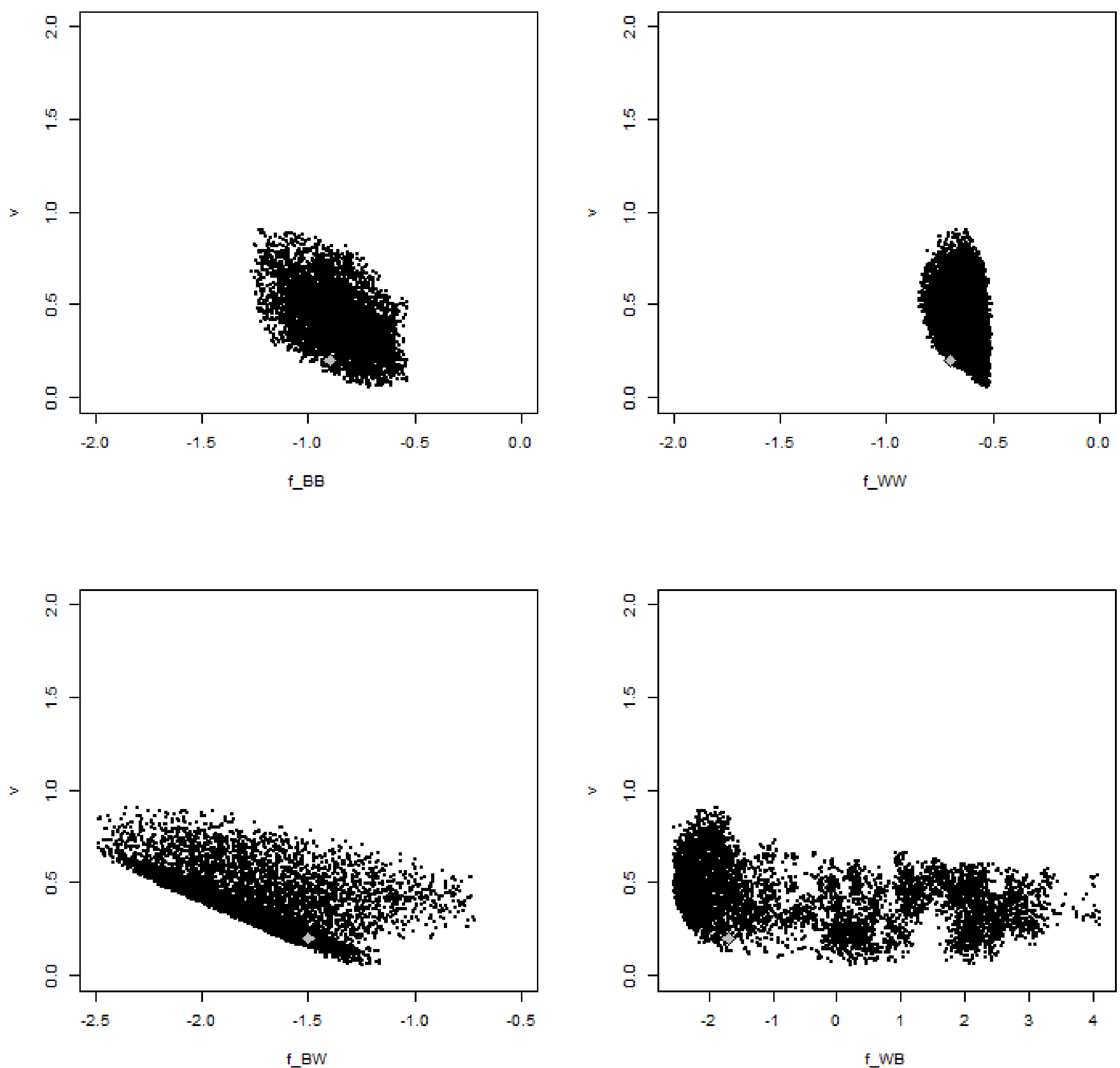
Notes: Figure plots shares of isolated types (x: isolated) and certain combinations of other types: types with any own-race friends (x: x friend(s)), any opposite-race friends (x: y friend(s)), any indirect friends, and any mutual friends. Points represent the shares from different simulated networks, and triangles indicate the shares from the network randomly selected to use as the observation.

Figure 10: Identified Set in Second Exercise: Values of Direct Connections



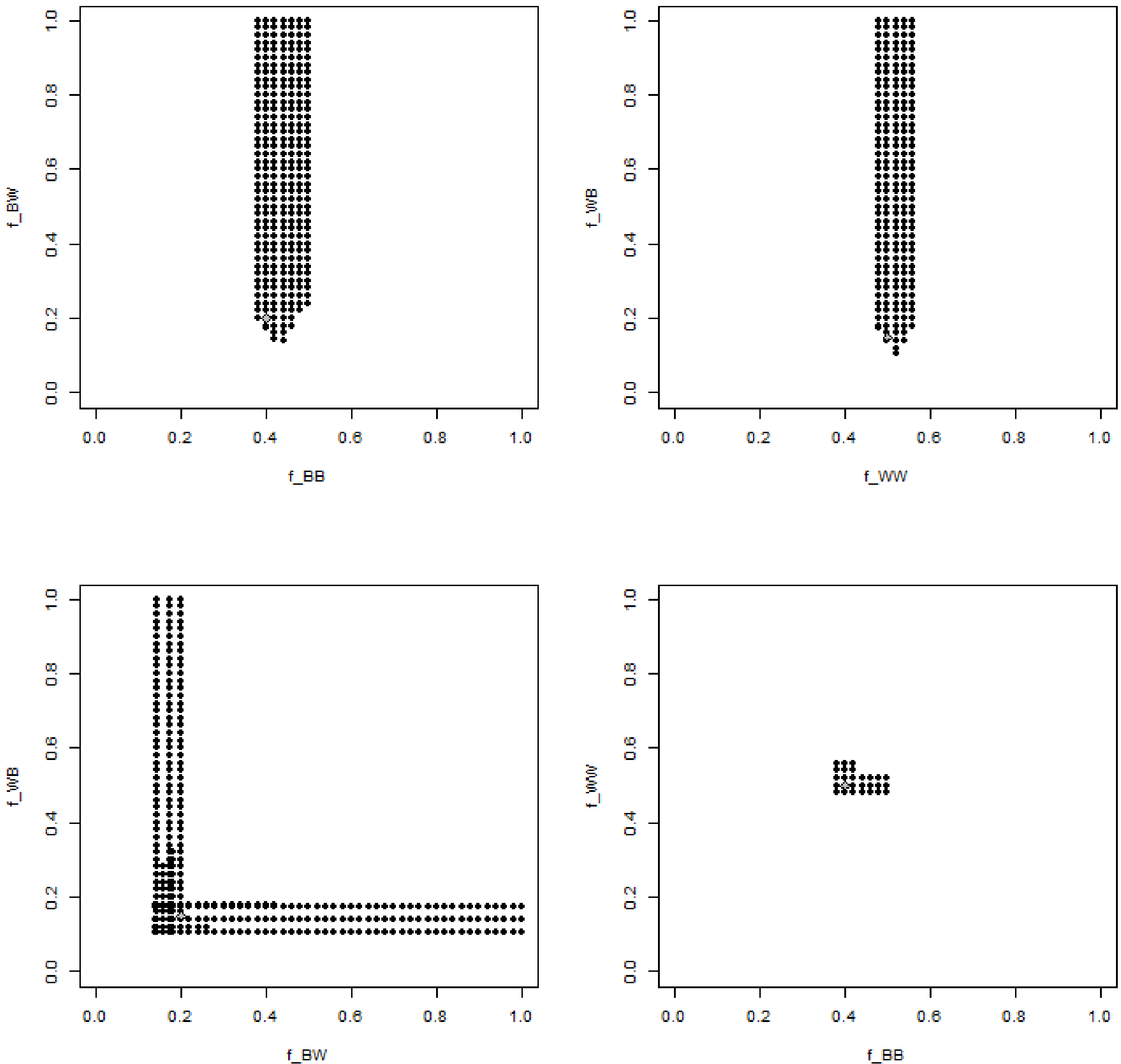
Note: Diamonds indicate true parameter values: $f_{BB} = -0.9$, $f_{BW} = -1.5$, $f_{WB} = -1.7$, $f_{WW} = -0.7$.

Figure 11: Identified Set in Second Exercise: Values of Indirect vs. Direct Connections



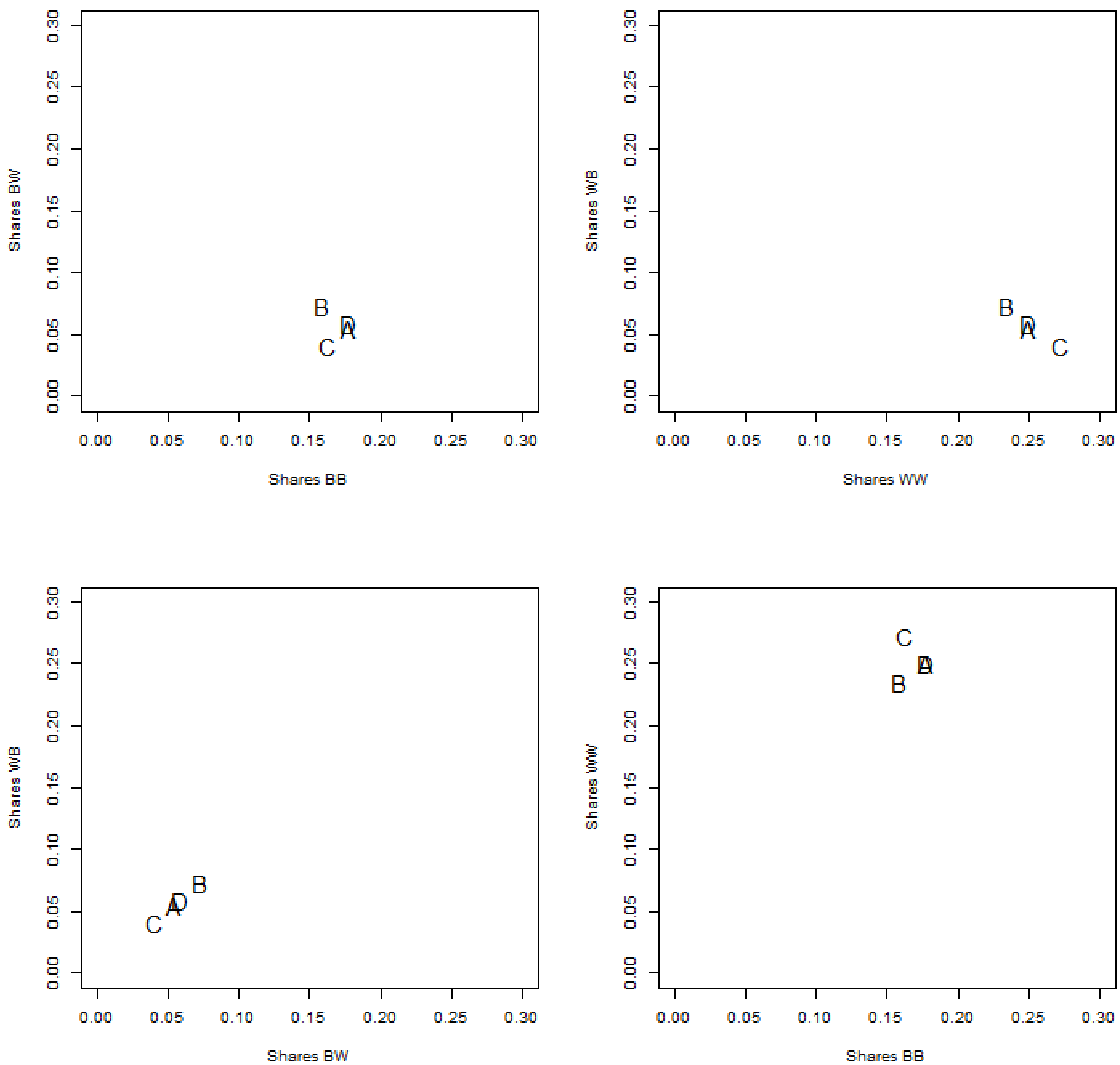
Note: Diamonds indicate true parameter values: $f_{BB} = -0.9$, $f_{BW} = -1.5$, $f_{WB} = -1.7$, $f_{WW} = -0.7$, and $\nu = 0.2$.

Figure A1: Identified Set from One Finite Network, Using Simplified QP Problem



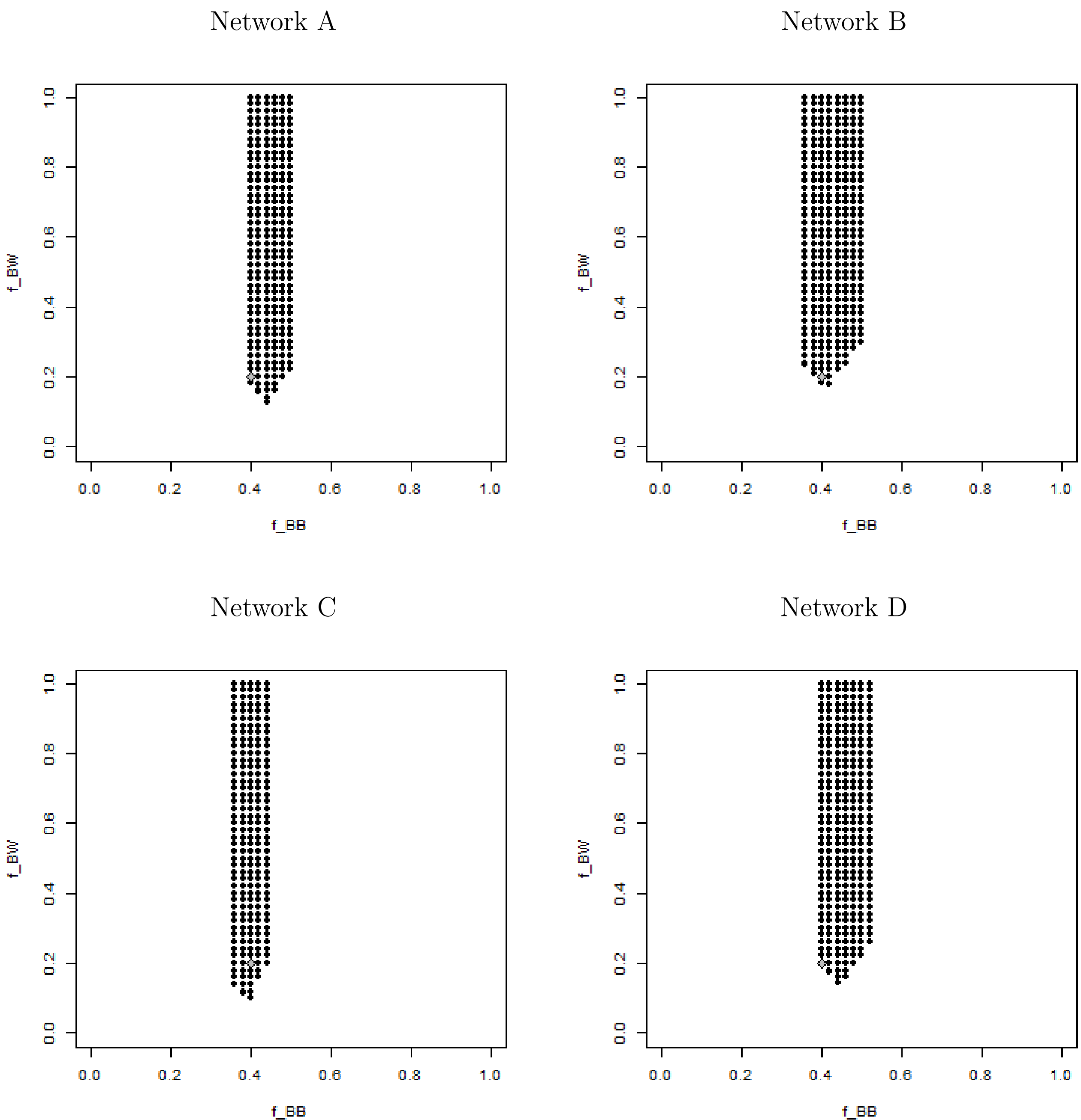
Note: Diamonds indicate true parameter values: $f_{BB} = 0.40$, $f_{BW} = 0.20$, $f_{WB} = 0.15$, and $f_{WW} = 0.50$.

Figure A2: Type Shares in Four Randomly Selected Networks (A, B, C, D)



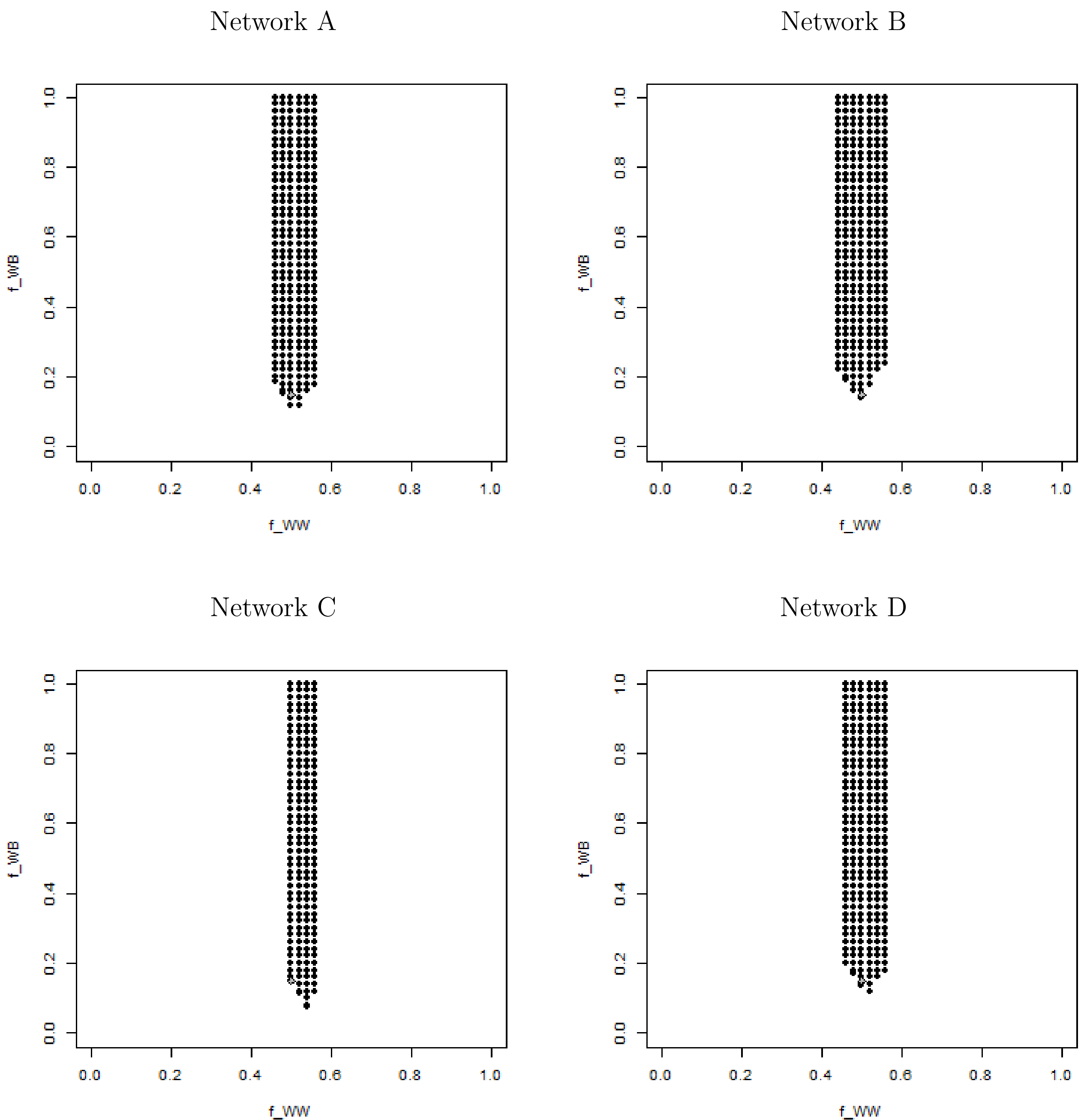
Notes: Letters A, B, C, D correspond to four vectors of type shares that were randomly selected from the full set of equilibrium type shares in Figure 6. Positions of the letters indicate the values of the type shares.

Figure A3: Identified Sets from Four Separate Networks: Black Preferences



Note: Diamonds indicate true parameter values: $f_{BB} = 0.40$, $f_{BW} = 0.20$.

Figure A4: Identified Sets from Four Separate Networks: White Preferences



Note: Diamonds indicate true parameter values: $f_{WW} = 0.50$, $f_{WB} = 0.15$.

Figure A5: Degree Distribution in Second Exercise Compared with Add Health Study