

de Paula, Áureo; Richards-Shubik, Seth; Tamer, Elie

Working Paper

Identification of preferences in network formation games

cemmap working paper, No. CWP29/15

Provided in Cooperation with:

Institute for Fiscal Studies (IFS), London

Suggested Citation: de Paula, Áureo; Richards-Shubik, Seth; Tamer, Elie (2015) : Identification of preferences in network formation games, cemmap working paper, No. CWP29/15, Centre for Microdata Methods and Practice (cemmap), London,
<https://doi.org/10.1920/wp.cem.2015.2915>

This Version is available at:

<https://hdl.handle.net/10419/130047>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Identification of preferences in network formation games

Áureo de Paula
Seth Richards-Shubik
Elie Tamer

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP29/15

Identification of Preferences in Network Formation Games*

Áureo de Paula[†]

UCL, São Paulo School of Economics,
CeMMAP, IFS and CEPR

Seth Richards-Shubik[‡]

Carnegie Mellon University
and NBER

Elie Tamer[§]

Harvard University

First Version: December 2011

This Version: June 2015

Abstract

This paper provides a framework for identifying preferences in a large network under the assumption of pairwise stability of network links. Network data present difficulties for identification, especially when links between nodes in a network can be interdependent: e.g., where indirect connections matter. Given a preference specification, we use the observed proportions of various possible payoff-relevant local network structures to learn about the underlying parameters. We show how one can map the observed proportions of these local structures to sets of parameters that are consistent with the model and the data. Our main result provides necessary conditions for parameters to belong to the identified set, and this result holds for a wide class of models. We also provide sufficient conditions—and hence a characterization of the identified set—for two empirically relevant classes of specifications. An interesting feature of our approach is the use of the economic model under pairwise stability as a vehicle for effective dimension reduction. The paper then provides a quadratic programming algorithm that can be used to construct the identified sets. This algorithm is illustrated with a pair of simulation exercises.

*We thank participants at the ASSA meetings in Chicago (2012), University of Western Ontario (2013) seminar, 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), the Heinz seminar, and the Cowles Foundation (2015) for suggestions and comments. We are also grateful to A. Advani, J. Altonji, C. Gualdani, P. Haile, and M. Shum for detailed comments. de Paula gratefully acknowledges financial support from the National Science Foundation through award SES-1123990, the European Research Council 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, CeMMAP, IFS and CEPR. E-mail: apaula@ucl.ac.uk

[‡]Carnegie Mellon University and NBER, E-mail: sethrs@cmu.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 a rich class of network formation models. Our model is based on 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 key defining feature of strategic network formation models is some form of externality that goes beyond direct links, such as the idea that friends of friends matter.

A network formation model that is founded on a well-defined preference structure for all the players is helpful in determining how outcomes develop given a particular policy or incentive system. Having an economically founded and estimable model thus becomes a useful ingredient in understanding the role of networks as mediators in determining final outcomes. Empirically sound network formation models can also be helpful in describing why certain networks emerge and not others. Additionally, with an estimable model, we are able to track the effects of various policies or frictions on the kinds and shapes of networks that may emerge.

The problem in analyzing data that are generated from such strategic models is that multiplicity of solutions and computational difficulties (especially when a large number of agents are involved) are pervasive. We suggest a computationally tractable approach that allows for multiple equilibrium networks to arise given a set of parameter values and covariate realizations. To achieve this we impose restrictions on the number of connections that might affect a person’s utility, as well as on the cardinality of their observable characteristics. This allows us to frame the problem in terms of a finite number of possible local connections and thereby reduce the dimensionality of the computational problem. In fact, the majority of the computation in our procedure is relatively easy to handle through a series of quadratic programming problems.

We suggest an empirical approach to the characterization of (set-)identified parameters that bypasses the selection of a particular equilibrium (when many are possible) and exploits directly the economic prediction under pairwise stability. Our approach is based on the key idea of “network types,” whereby each individual is classified into one of a finite set of observed, mutually exclusive and exhaustive, local subnetworks that arise in equilibrium.

The set of relevant types is determined by the specification of the preference structure, so for example a specification where only direct connections matter will suggest a different set of types than a specification where indirect connections matter. The link between the *observed* frequencies of network types and their model predicted ones allows us to learn about preferences and preference parameters. Developing this correspondence in a computationally tractable way represents the main contribution of this paper. This allows us to construct sets of preference parameters that are consistent with observable data. The main result in the paper 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. Also, for certain classes of empirically relevant preference structures, we show that these conditions are sufficient as well, and so are able to characterize the identified set.

Related Literature

In part because of the difficulties indicated above, the literature on the econometric analysis of network formation models is small, but growing. We focus on complete information models with undirected links and non-transferable utility, and our solution concept is pairwise stability (Jackson and Wolinsky 1996). The payoff structures we analyze are related to those contemplated in, for example, Currarini, Jackson, and Pin (2009), Christakis, Fowler, Imbens, and Kallianaraman (2010), Sheng (2014), and (for a directed network) Mele (2013). 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 2014, Leung 2015).¹ Relative to these papers, our proposed method does not require certain categorical 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 Kallianaraman 2010, Mele 2013, Badev 2013, Chandrasekhar and Jackson 2014).^{2,3} 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 2014, Dzanski 2014).

¹Boucher and Mourifié (2013) develop a method that is similar to Leung (2015) but assumes uniqueness.

²These models are intended for use with static observations of network links, however, because dynamic data on link formation are rarely available. As pointed out in Mele (2013), the meeting protocol therefore acts as an equilibrium selection mechanism.

³Additionally, Badev (2013) extends the model to include actions beyond link formation, and Chandrasekhar and Jackson (2014) contain other interesting approaches to estimation using extensions of exponential random graph models.

Because matching models essentially aim at characterizing a bipartite graph, and hence a particular type of network, those models are also related to the literature on strategic network formation. There is a growing literature on the econometrics of matching models (e.g., Choo and Siow (2006), Fox (2010a), Fox (2010b), Galichon and Salanie (2009), Echenique, Lee, and Shum (2010), Chiappori, Galichon, and Salanié (2012), Menzel (2014b)). Our setting differs from those in substantive aspects: indirect connections are payoff relevant, and multiple equilibria are possible (in contrast to some, though not all, papers in that literature). Also the concept of pairwise stability in matching games is related, but not identical to the one in Jackson and Wolinsky (1996) where only one link at a time is considered.

2 Model Specification and Solution Concept

Our framework considers complete information games that produce an undirected network. One common 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 depart from the standard discrete choice literature in that we use a continuum of players $i \in N \equiv [0, \mu]$, where $\mu > 0$ is their total measure. This modeling choice is natural in our setting where we are considering games with a large number of players (Khan and Sun 2002). Each player has some predetermined characteristic(s) $X_i \in \mathcal{X}$ observed by the econometrician, and player-pairs have a one-dimensional characteristic ϵ_{ij} which 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. We assume that \mathbf{X} and ϵ are independent random vectors (though this is not essential to the analysis).

The network that results from players' actions is characterized by the adjacency mapping

$$G : N \times N \rightarrow \{0, 1\}.$$

This is a *continuous graph* as there is a continuum of nodes.⁵ Such graphs (particularly refinements known as *graphons*, for limits of dense graphs, or *graphings*, for limits of bounded degree graphs) are a recent development (for a review, see Lovasz (2012)) and are used as

⁴Pairwise stability can also define equilibrium in dynamic games, as a rest point of a sequential meeting process (Jackson (2009)).

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

approximations for large graphs under a well-defined metric. Hence we view the continuous graph model here as a close approximation to a model with a large (but finite) number of players.

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 provide an approach to learn about the (parameterized) payoff functions $u(G, \mathbf{X}, \cdot)$ using the data. 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 are utility (or payoff) relevant, and preferences are such that players will never choose more than a total of L links.*

In the above, the distance between two agents refers to the length of the shortest path between those individuals, denoted $d(i, j; G)$. If the distance between two individuals is finite, but there is not a direct link between them, we say these two agents are *indirectly connected*. By Assumption 1, if $D = 1$, one does not have a taste for who your indirect connections are. When indirect connections do matter, most specifications in the literature appear to rely on $D = 2$. On the other hand, L denotes the maximum number of links an individual would have (utility would be infinitely negative if you have more than L links).⁶ Together, these restrictions make it so that payoffs depend on a finite number of direct and indirect connections in the network. For example with $D = 2$, there would be at most L direct alters and $L \times (L - 1)$ indirect alters that impact utility. This assumption also leads to relatively sparse equilibrium networks, since the number of links per person is small relative to the total number of possible connections. This sparsity is empirically plausible in many networks observed in the social sciences (see, e.g., Backstrom, Boldi, Rosa, Ugander, and Vigna (2012)).

With a finite number of links and finite support for the predetermined characteristics X (see Assumption 2 below), there is a finite number of possible configurations of alters

⁶For 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.

and their characteristics up to any finite distance from a given individual. Our proposed inference strategy relies on this feature of the local structure around each individual in the equilibrium network to reduce the dimensionality of the problem from the universe of possible network configurations to the possible categories of payoff-relevant local subnetworks.⁷ The limitations on size of the relevant local subnetwork from Assumption 1, and the finite support for X , indicate that a finite number of possible local structures will play a key role in the identification of preference parameters.⁸

A second assumption we make on the payoff structure is that the preference shocks do not depend on the individual identities of the alters. Similar anonymity restrictions are common in models of large games (Kalai (2004), Menzel (2014a), Song (2014)). Instead, we assume that there are as many shocks as the maximum number of direct connections one might establish, crossed with their possible predetermined characteristics. The number of preference shocks for each individual thus equals $L \times |\mathcal{X}|$, where $|\mathcal{X}|$ is the cardinality of \mathcal{X} . This assumption helps us control the dimensionality of the problem. Furthermore, this is a natural restriction to make in models with large numbers of players, where individual identities are unknown. It also allows the model to retain a positive fraction of isolated individuals in equilibrium even when the group under consideration is large. 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). In our model, even with large networks, it is still possible to have isolated nodes. Thus we have our second assumption:

Assumption 2. *There is one preference shock $\epsilon_{il}(x)$ for each potential direct connection $l = 1, \dots, L$ and each characteristic $x \in \mathcal{X}$. This vector of preference shocks is independent of X with a known distribution (up to some finite dimensional parameter). In addition, the support of X (i.e., $\text{supp}(\mathcal{X})$) is finite.*

⁷The use of subnetworks here is rather distinct from the use in Sheng (2014). 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.

⁸ In principle, any set of finitely supported random vectors would be eligible for the predetermined characteristics in the algorithm we suggest. However one may end up with a large number of possible local network categories even in comparison to the number of individuals in the network. In this case, not to add to the computational complexity of our suggested methodology, we recommend that individuals be clustered into a small number of groups in the initial stages of the analysis according to their observable characteristics. Well known algorithms for data segmentation are available in the statistical literature (e.g., Chapter 14 in Hastie, Tibshirani, and Friedman (2001)).

Remark 1. *This assumption could be relaxed to allow for shocks to each potential payoff-relevant connection (both direct and indirect), and our main results would remain unchanged. In addition, we can allow the vector of preference shocks to be correlated. This has economic relevance since it captures preference correlation among unobservables. However, generally allowing for unrestricted correlation might lead to exceedingly large identified sets.*

2.1 Payoff Examples

Next we showcase two classes of preference specifications that are used in the literature.

Example 1 (friendship game with “same” and “different” friends). *Consider an empirical version of the game in Currarini, Jackson, and Pin (2009), where the utility function is*

$$u_i(G, \mathbf{X}) = u(x_i, m, n; \epsilon_i),$$

in which m is the number of direct connections (e.g., friends) with the same predetermined characteristic as i and n is the number of direct connections with a different characteristic. A person is assumed to establish (at most) L connections, and $D = 1$ so there is no taste for indirect connections. As in that paper’s application, the predetermined characteristics X are races. The vector of preference shocks ϵ_i contains one shock for each potential connection to a same-race individual and one shock for each potential connection to a different-race individual, for a total of $2 \times L$ elements (not $|\mathcal{X}| \times L$ elements). In other words, here the dimension of the preference shocks is reduced further by assuming that the shocks for links to people with any different characteristics are perfectly correlated. We accordingly denote the elements of ϵ_i as $\epsilon_{il}(z)$, where the index $l = 1, \dots, L$ refers to an enumeration of the potential connections and $z = 0, 1$ indicates whether the alter is of same ($z = 0$) or different ($z = 1$) race; thus, $\epsilon_i \equiv (\epsilon_{i1}(0) \dots \epsilon_{iL}(0), \epsilon_{i1}(1) \dots \epsilon_{iL}(1))$. In addition, let the specification of the utility function be as follows:

$$u(x, m, n; \epsilon_i) \equiv (m + \gamma n)^\sigma - \sum_{l \leq m} \epsilon_{il}(0) - \sum_{l \leq n} \epsilon_{il}(1).$$

This says that if one has m same-race friends and n different-race friends, one subtracts the sums of the first m shocks for same-race alters and the first n shocks for different-race alters. We could think of these shocks as expressing the costs of maintaining these friendships. The term $(m + \gamma n)^\sigma$ is taken from Currarini, Jackson, and Pin (2009). The $\sigma \in (0, 1)$ gives diminishing returns. If $\gamma < 1$ there is a preference for same-race friendships, while $\gamma > 1$ would express a preference for diversity.

Example 2 (friendship game where indirect connections matter). *Suppose individuals can form up to L direct friendships. For each one of those, let $j(l)$, $l = 1, \dots, L$, denote the person in N with whom that connection is established. (Set $j(l) = \emptyset$ if the l^{th} friendship is not established.) Utility functions in this class are given by:*

$$\begin{aligned}
u_i(G, \mathbf{X}) \equiv & \sum_{\substack{l=1 \\ j(l) \neq \emptyset}}^L G(i, j(l)) (f(x_i, x_{j(l)}) + \epsilon_{il}(x_{j(l)})) & (\text{direct connections}) \\
& + \left| \bigcup_{l=1, j(l) \neq \emptyset}^L N(j(l)) - N(i) - \{i\} \right| \nu & (\text{friends of friends}) \\
& + \sum_{\substack{l=1 \\ j(l) \neq \emptyset}}^L \sum_{\substack{k>l \\ j(k) \neq \emptyset}}^L G(j(l), j(k)) \omega & (\text{mutual friends})
\end{aligned}$$

where $N(i)$ denotes the set of nodes directly connected to node i and $|\cdot|$ is the cardinality of a given set. Notice that, in this example, $D = 2$. Variations of this specification have been used widely in the literature (e.g., Christakis, Fowler, Imbens, and Kalianaraman (2010), Goldsmith-Pinkham and Imbens (2013), Sheng (2014)).⁹ Notice that the connections here are not double counted. Double counting and the positivity of ν and ω are sometimes used to establish the existence of a pairwise stable network. We do not need to impose this because, in our case, when the data cannot be reconciled with equilibrium play for a given model, our procedure can deliver an empty identified set.

2.2 Solution Concept

As in the empirical games literature, we will assume that observed choices correspond to equilibrium play. Next, we define the equilibrium solution concept of pairwise stability that we use.

Definition 1 (Pairwise Stability, Jackson and Wolinsky (1996)). *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:*

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

$$\forall i, j : G(i, j) = 0, \quad \text{if } u_i(G_{+ij}, \mathbf{X}) > u_i(G, \mathbf{X}) \text{ then } u_j(G_{+ij}, \mathbf{X}) < u_j(G, \mathbf{X}). \quad (2)$$

⁹Christakis, Fowler, Imbens, and Kalianaraman (2010) further allow for a quadratic term on the number of friends of friends and an additional term for multiple paths to an *indirect* friend.

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 $(k, l) = (i, j)$. An implication of (2) is that the addition of absent links to a pairwise stable network would decrease the utility of at least one of the two people involved. As noted in the literature, this concept is different from the Gale-Shapley notion of stability in two-sided (e.g., “marriage”) matching games where two couples cannot be made better off by recombination. Here the evaluation of stability is performed one link at a time.

Other solution concepts exist, such as the Nash Equilibria to a links-announcement game; the intersection of the set of Nash Equilibria and pairwise stable networks; and a subset of those equilibria where deviations by coalitions are considered (i.e., strongly stable networks). For more on this, 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. Whereas refinements of pairwise stability exist, we view this notion as defining intuitive necessary conditions for equilibrium in many settings.

2.3 Preview of Our Results

Below, we develop a set of aggregate conditions that tie the distribution of preferences in the population to the frequencies of *network types* (the local network structures alluded to earlier, formally defined below in Section 3) that would be observed in an equilibrium network. We show that these conditions are necessary for a pairwise stable network to exist with given frequencies of each network type (i.e., data) for payoff structures that satisfy Assumptions 1 and 2. We also show that these conditions are both necessary and sufficient for the models in Examples 1 and 2 (subject to a restriction in the latter case). Hence for these empirically relevant classes of models we are able to obtain sharp bounds on the identified set of preference parameters.

In the remainder of this section we provide a preview of our approach and these results using a simple case of Example 1. The main role of this example is to illustrate key features of the approach as simply as possible. We will return to it in later sections to clarify the concepts introduced here and formalized throughout the paper. The illustrative model we

use is one of choosing best friends among individuals of two races: B (black) or W (white). Only direct connections affect utility ($D = 1$) and individuals have at most one best friend ($L = 1$).

Payoffs are a function of the individual's race, X_i , the best friend's race, X_j (where $X_j = 0$ if i is isolated), and the individual's preference shocks, ϵ_i . The shocks depend on the friend's race, but not on her identity (Assumption 2). This and the finite cardinality of the set of characteristics \mathcal{X} ($\{B, W\}$ in this example) play an important role in reducing the dimensionality of the problem. Instead of having one preference shock for each potential mate, each person draws only two shocks, one for each race of the potential best friend: $\epsilon_i \equiv (\epsilon_i(B), \epsilon_i(W))$. The utility function is specified as

$$u(X_i, X_j, \epsilon_i) = f_{X_i X_j} + \epsilon_i(X_j) \quad (3)$$

where $(f_{xy})_{x,y \in \mathcal{X}}$ is the parameter of interest that expresses the deterministic component of the payoff to an individual of race x from having a best friend of race y . The utility of being isolated is normalized to zero.

Equilibrium outcomes in this simple network can be expressed as an ordered pair (x, y) for the individual's race and the best friend's race (or $y = 0$ for no best friend). For example, (B, B) corresponds to a black individual with a black best friend (in equilibrium). These pairs represent the **network types** in this simple model. Each individual is one network type, and we observe the equilibrium proportions of these types. The objective is to use these proportions to learn about $(f_{xy})_{x,y \in \mathcal{X}}$.

To link the proportions to the preference parameters, we start by classifying individuals based on which network types have links they would be happy to keep. For example, depending on the preference shocks drawn, a black individual may prefer having a black best friend to being alone, but may prefer being alone to having a white best friend. Hence the network type with a white best friend (B, W) would not be an equilibrium outcome for this individual, but (B, B) could be. We refer to the sets of network types that individuals would not unilaterally break away from as **preference classes**. (Because there are no connections to be dropped from an isolated type, e.g. $(B, 0)$, these network types belong to all preference classes.) Heuristically, a preference class can be identified with a region in the space of the shocks (ϵ) that determines the network types from which a person would not prefer to unilaterally drop a link. Given the preference shock distribution and proposed values for the preference parameters, one can then compute the probability that an individual with given

characteristics pertains to a particular preference class.

We characterize a pairwise stable network by allocating the individuals in each preference class to the possible network types. In this example, there are eight preference classes (four for each individual race) and six network types (three for each race). Consequently, one can define $6 \times 8 = 48$ **allocation parameters** indicating the proportion of agents in each preference class allocated to each of the observable network types. Then, given a vector of structural parameters and distribution of preference shocks, which determine the probability of each preference class, we can use the allocation parameters to obtain the proportions of individuals of each network type.

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. In fact, we show that these restrictions amount to finding allocation parameters that minimize a well-defined quadratic objective function at zero given constraints. The quadratic objective and these constraints correspond to necessary equilibrium restrictions and agreement with the data. First, individuals may only be allocated to network types admissible to 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 these 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 proportions of types obtained from the allocation parameters must match the observed proportions of types in the network.

Hence, given a vector of structural preference parameters, if we are able to find allocation parameters that solve this quadratic programming problem attaining a zero objective function value, the parameter vector belongs to the identified set. Under certain conditions on preferences and distributions of shocks we are further able to establish that the restrictions above are not only necessary but also sufficient, and so the set formed by collecting the structural parameters that lead to a zero solution is the sharp identified set.

3 Network Types and Preference Classes

We now formalize the objects introduced in the above example. As we have noted earlier, the main reason for taking this approach is the computational burden that would otherwise be involved in searching for equilibrium networks, given a candidate vector of preference parameters. For a finite group of n individuals there are $2^{\frac{n(n-1)}{2}}$ possible networks. In a typical network with 100 individuals, for example, this would correspond to 2^{4950} ($\approx 10^{1500}$) possible configurations. For comparison, this is considerably larger than the estimated number of atoms in the universe. Since we aim for a framework to study a potentially large network, methods designed in the empirical games literature (on entry for example) are simply infeasible due to this computational burden.

This paper instead proposes a framework for studying large networks that relies on a particular approach to dimension reduction via anonymity and aggregation. Given the restriction on preferences (only L links are allowed, and only depth D matters), utility is fully determined by an individual’s *network type*, more formally defined below. A utility structure generates a set of payoff-relevant network types, and a parameterization of the continuous graph model of network formation 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.

We now turn to the definition of network types. *Our proposed identification strategy relies on pre-defined network types that describe the local network structure around an individual along with the predetermined characteristics of the individuals in this local subnetwork, and whose estimated proportions in the sample are an equilibrium outcome.* The predetermined characteristics are observed 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.¹⁰ A network type can then be intuitively described, for instance, “a female connected to two females and no males,” “an unconnected low-income male,” “a female connected to another female with at least two friends,” etc. Each person in the data belongs to one of these mutually exclusive and exhaustive types or categories, and we assume that the proportions of individuals of each network type can be consistently estimated from the observed data. (Additionally, it may be possible to estimate the proportion of network types even without complete observation of the network—see, e.g., Kolaczyk (2009).)

¹⁰As indicated in footnote 8, one can possibly cluster individuals into a small number of groups in the initial stages of the analysis according to observable characteristics using well-known algorithms for data segmentation in the statistical literature.

More precisely, given D and L , a network type describes the local subnetwork up to distance D from the reference individual, who is called the *ego* of the network type. Network types are thus characterized by a local adjacency matrix, A , and a column vector of the predetermined characteristics of the ego and alters in the subnetwork, v . The matrix A is square and has one row (and column) for the ego and one for each alter up to depth D , for a total of $1 + L + L(L-1) + L(L-1)^2 + \dots + L(L-1)^{D-1} = 1 + L \sum_{d=1}^D (L-1)^{d-1}$ rows.¹¹ Since A represents a simple undirected graph, it is symmetric with zeros on the diagonal. Note that the dimension of this adjacency matrix does not depend on the size of the network. The number of rows in A is also the number of elements in the vector v . The first element of v is the characteristic of the ego, denoted v_1 . The subsequent elements are the characteristics of the alters, from $\mathcal{X} \cup \{0\}$, where 0 denotes the absence of an alter in that position. We summarize this definition below.¹²

Definition 2. *Fix D , L , and \mathcal{X} . Each person belongs to a **network type** t . 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 the set \mathcal{T} . The network type of each individual is determined in equilibrium and is observed in the data.*

¹¹This 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).

¹²Each network type is actually a collection of isomorphic matrix and vector pairs (A, v) . Two graphs with adjacency matrices G_1 and G_2 are isomorphic if and only if there exists a permutation matrix P such that $P^\top G_1 P = G_2$. The first row and first 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. Thus there will be more than one pair (A, v) corresponding to the same network type, and the elements in this equivalence class are obtained by permutation of friends. Finding whether or not two graphs are isomorphic has an unknown computational complexity. (It is known to be in **NP**, but it is not known if it is in **NP**-complete, or if it is in **P**.) For computational convenience, then, one should adopt a convention to single out a representative element from each class. We adopt the convention that after the first line (row/column of A or element of v), which corresponds to the individual of interest (the ego), the subsequent L positions correspond to her direct connections (or direct alters). Then the next $L-1$ lines correspond to the $L-1$ additional possible links of the first direct alter, and so on. Should the ego have fewer than L links, we position the 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 applies to any alter that does not have his full set of links as well. 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} , to be adopted in aligning the characteristics in v , would fix the permutation and provide a canonical element from the equivalence class.

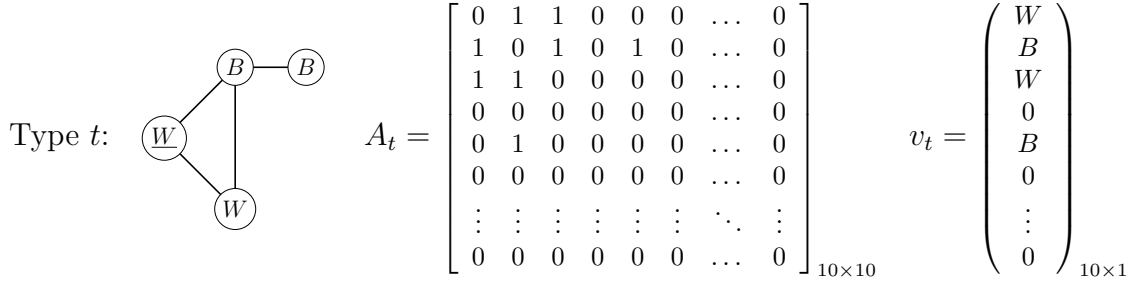


Figure 1: Illustration of a network type, with matrix A and vector v .

Example 3. This picks up on the simple model from Section 2.1 where $D = L = 1$ and $\mathcal{X} = \{B, W\}$. There are two possible local adjacency matrices (which are 2×2): A_0 (unlinked) and A_1 (one link). There are six vectors of characteristics: $v \in \{B, W\} \times \{0, B, W\}$. There are then six network types: $A = A_0$ and $v = (B, 0)$ or $(W, 0)$; $A = A_1$ and $v = (B, B)$, (B, W) , (W, B) , or (W, W) . (Note that A is redundant here so we could just use v in this example.) The data provide the frequency of the three network types for each race: the proportions of “black alone,” “black with black friend,” and “black with white friend” for blacks, and similarly for whites. Note that these types are derived from the specification of preferences and so there is a sense in which network types are all that the model “predicts” in terms of observables.

Example 4. Next, we consider an example from the class of models where indirect connections matter (Example 2). The predetermined characteristic is race, $\mathcal{X} = \{W, B\}$, and $D = 2, L = 3$; i.e., the maximum number of direct links is 3, and only indirect connections up to depth 2 are utility relevant. Then, a network type would specify the network structure up to depth 2 and the race of each alter in this subnetwork. For example, the type t illustrated in Figure 1 is a white individual (the characteristic of the ego is underlined) with one black friend and one white friend, who are also friends with each other, and where the black friend has a further friend who is black. This graph can be equivalently represented by the local adjacency matrix and vector of characteristics (A_t, v_t) as shown. In the utility specification for this class of models, the utility of this type for individual i would be $(f(W, B) + \epsilon_{i1}(B)) + (f(W, W) + \epsilon_{i2}(W)) + \nu + \omega$.

Next we show how utility is expressed as a function of an individual’s network type. From our first assumption, utility depends only on the local subnetwork that corresponds to

a network type. Hence we can replace G and \mathbf{X} in the utility function with A and v . From our second assumption, the preference shocks depend on the positions of the direct alters within the network type and their predetermined characteristics. The position of an alter corresponds to some row \mathbf{k} in the matrix A , and the characteristic of that alter would be $v_{\mathbf{k}}$. The number of shocks for each individual equals the number of potential direct connections multiplied by the cardinality of \mathcal{X} (i.e., $L \times |\mathcal{X}|$). If, instead, shocks to every potential payoff-relevant connection were allowed for, the number of shocks would be $|\mathcal{X}| \times L \sum_{d=1}^D (L-1)^{d-1}$ (for $L > 1$). Given these assumptions, we do not need any information beyond A , v , and ϵ_i to determine the payoffs for player i . Hence the utility function can be defined as

$$u_i(A, v) = u(A, v; \epsilon_i).$$

Apart from Assumptions 1 and 2, no further restrictions on the preferences are imposed. In particular, this formulation allows for negative externalities. Whereas this may pose problems for existence of a pairwise stable network (i.e., with negative externalities, it is easy to show that pairwise stable networks may not exist), non-existence in our setup is not problematic, as our methodology could then deliver an empty set for the identified set of parameters.

Remark 2. *It is worth emphasizing here that the set of network types is determined completely by the specification of preferences (and the set of predetermined characteristics X). 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.” By considering all of the possible types that are relevant under a particular class of preferences, there is no loss of information to make inferences about these preferences.*

Important in our characterization of stability is what we call a *preference class*. As we pointed out in Section 2.1, these preference classes are sets of network types that categorize individuals based on which types correspond to links they would be happy to keep. Given an individual's preference shocks (ϵ_i), a preference class is the set of network types that would satisfy *that individual's* stability requirement for existing links. This corresponds to expression (1) in the definition of pairwise stability, from the perspective of one of the two individuals involved in a given link. In order to state this requirement using a utility function defined on (A, v) rather than on (G, \mathbf{X}) , we first define the matrix A_{-l} to be equal to A but

with the l th direct link removed.¹³ Then we can say the pairwise stability condition in (1) requires that

$$\forall i \in N, \forall l = 1, \dots, L, \quad u(A, v; \epsilon_i) \geq u(A_{-l}, v; \epsilon_i).$$

An individual represented by X_i and ϵ_i can then be classified as belonging to the preference class comprised of those types (A, v) with $v_1 = X_i$ such that the above condition is satisfied. If assigned to any of those network types, this individual is content to keep his or her connections, whereas he or she would be inclined to drop connections when allocated to types outside the preference class. We formalize the definition below.

Definition 3. Let $(x, \epsilon) \in \mathcal{X} \times \text{supp}(\epsilon) \mapsto H(x, \epsilon) \subseteq \mathcal{T}$ that assigns to each pair (x, ϵ) the network types that satisfy $u(A, v; \epsilon) \geq u(A_{-l}, v; \epsilon)$ where $v_1 = x$ and $l = 1, \dots, L$. The set of types $H \subseteq \mathcal{T}$, or **preference class**, is then an element in the range of the mapping $H(\cdot, \cdot)$.

Note that all preference classes include an isolated type. There are no links to drop from these types, so $A = A_{-l}$. It is also helpful to note that, given a particular realization of X , preference classes partition the ϵ space. Those individuals whose preference shocks fall within a given region of this partition would not want to drop any connections if assigned to one of the network types in the corresponding preference class, but would do so if assigned to any network type outside of that preference class. Finally, to understand the construction of preference classes in the definition above, consider for example $H = \{t_1, t_2\}$ for some $t_1, t_2 \in \mathcal{T}$. Then H is indeed a preference class if, for a realization of X with the relevant value, there is a region for ϵ such that an individual with those features would not want to drop a link if allocated to network type t_1 or t_2 , but would do so if allocated to any other network type. This is illustrated in the example below.

Example (3 cont'd). As indicated earlier, network types in this simple model can be expressed as an ordered pair $(x, y) \in \mathcal{X} \times (\mathcal{X} \cup \{0\})$ for the individual's race and the best friend's race (or $y = 0$ for no best friend). We now describe the preference classes. For blacks there are four possible sets of types that could comprise a preference class depending on the values of $\epsilon_i(B)$ and $\epsilon_i(W)$:

¹³In other words, set elements $(1, l+1)$ and $(l+1, 1)$ in A to zero. There may remain nodes in the subnetwork represented by A_{-l} that are irrelevant to the 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 $u(A_{-l}, v; \epsilon_i)$.

- $H_1 = \{(B, 0)\}$: a friend of neither race would be acceptable;
- $H_2 = \{(B, 0), (B, B)\}$: only a black friend would be acceptable, otherwise alone;
- $H_3 = \{(B, 0), (B, W)\}$: only a white friend would be acceptable, otherwise alone; and
- $H_4 = \{(B, 0), (B, B), (B, W)\}$: either race would be acceptable, otherwise alone.

Note that these preference classes partition $\text{supp}(\epsilon) \in \mathbb{R}^2$ into four regions: $\{\epsilon \in \text{supp}(\epsilon) : \epsilon_B < -f_{BB} \text{ and } \epsilon_W < -f_{BW}\}$ for the first preference class; $\{\epsilon \in \text{supp}(\epsilon) : \epsilon_B \geq -f_{BB} \text{ and } \epsilon_W < -f_{BW}\}$ for the second; $\{\epsilon \in \text{supp}(\epsilon) : \epsilon_B < -f_{BB} \text{ and } \epsilon_W \geq -f_{BW}\}$ for the third class; and $\{\epsilon \in \text{supp}(\epsilon) : \epsilon_B \geq -f_{BB} \text{ and } \epsilon_W \geq -f_{BW}\}$ for the fourth class.

Also, a given type can belong to many preference classes as is clear from the example above, where “black alone” (the isolated type) belongs to all preference classes for blacks, and type (B, B) belongs to the second and the fourth preference classes for blacks.

The probability of some preference class H is the probability of the region in ϵ that would support all the types in H . In other words, all the types in H would satisfy $u(A, v; \epsilon) \geq u(A_{-l}, v; \epsilon)$, $l = 1 \dots L$, for any value of ϵ in this region (and no other types would). Because all the network types in a given preference class must have the same predetermined characteristic of the ego, v_1 , we define these probabilities conditionally on that characteristic:

$$P_{H|v_1} \equiv P(\epsilon : H(X, \epsilon) = H | X = v_1).$$

These probabilities are direct functions of the utility specification and the distributions of the unobservables. Hence, these probabilities are known given a parameterization of the model.

Example (3 cont’d). An individual i would prefer being best friends with someone of race y over being isolated if and only if $f_{X_i y} + \epsilon_i(y) \geq 0$. So the probability that an individual of race x would prefer being best friends with someone of race y over being isolated is

$$p_{xy} \equiv 1 - F_\epsilon(-f_{xy}) \tag{4}$$

where F_ϵ is the c.d.f. of $\epsilon(y)$, for $y = B, W$. If $\epsilon(B)$ and $\epsilon(W)$ are i.i.d., the probability of each preference class is easily expressed in terms of the p_{xy} defined in (4). For $H = \{(B, 0), (B, B)\}$, for example, the probability that these outcomes would satisfy the preference

condition for a black individual is $p_{BB}(1 - p_{BW})$. This is the probability that the individual would be content to establish a link with a black person but not with a white person.

Finally, we can link the model to the data using the estimable proportions of individuals in the network who are of each network type. To generate predictions from the model we specify how many individuals from each preference class are assigned to the various network types. Accordingly we define *allocation parameters*, which give the proportions of individuals in each preference class that are assigned to each network type.

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

$$\mu_{v_1} \sum_H P_{H|v_1} \alpha_H(t),$$

where μ_{v_1} is the measure of individuals with characteristic v_1 . The proportion of individuals of this type in the network is the above measure divided by the total measure of individuals in the network, μ .

The measures (or proportions) of network types in the definition above provide the exact link between the data and the underlying preferences. The proportions of individuals of each network type can be consistently estimated, and we use the predictions from the model to learn which preference parameters could be consistent with the observed data. The example below illustrates how this works, in general terms.

Example (3 cont'd). As noted earlier, with i.i.d. preference shocks the conditional probabilities of each preference class can be stated as, for example, $p_{BB}(1 - p_{BW})$ for class $\{(B, 0), (B, B)\}$. We use these probabilities and the allocation parameters to determine the proportions of individuals of each network type. For example the measure of blacks with white best friends (type (B, W)) would be as follows (multiplied by μ_B):

$$\begin{aligned} & \alpha_{H_1}(B, W) \underbrace{(1 - p_{BB})(1 - p_{BW})}_{P_{\{(B,0),(B,B)\}|B}} + \alpha_{H_2}(B, W) \underbrace{p_{BB}(1 - p_{BW})}_{P_{\{(B,0),(B,B)\}|B}} \\ & + \alpha_{H_3}(B, W) \underbrace{(1 - p_{BB})p_{BW}}_{P_{\{(B,0),(B,W)\}|B}} + \alpha_{H_4}(B, W) \underbrace{p_{BB}p_{BW}}_{P_{\{(B,0),(B,B),(B,W)\}|B}} \end{aligned}$$

where, as previously defined, $H_1 = \{(B, 0)\}$, $H_2 = \{(B, 0), (B, B)\}$, $H_3 = \{(B, 0), (B, W)\}$, and $H_4 = \{(B, 0), (B, B), (B, W)\}$. Because (B, W) is not in the first two preference classes, those allocation parameters must be zero in equilibrium. Hence, once pairwise stability is imposed, the proportion of blacks of network type (B, W) is effectively

$$\alpha_{H_3}(B, W)(1 - p_{BB})p_{BW} + \alpha_{H_4}(B, W)p_{BB}p_{BW}$$

This must equal the observed proportion of blacks with white best friends.

Conditions such as these, obtained from equilibrium requirements, will provide the link between the preference parameters, which determine the probabilities of each preference class, and the observed proportions of network types. In order to be included in the identified set, a vector of preference parameters must be able to generate the proportions of types observed in the data while satisfying such conditions on the allocation parameters.

4 Identification with Network Types

As noted above, the data allow us to learn the proportions of individuals of each type in a network (or more compactly, the *type shares*). In this section, we formalize how to use the model and its restrictions via pairwise stability to map the observed type shares into restrictions on the preference parameters. First, a set of general conditions is developed using necessary conditions for pairwise stability, which collects preference parameters that could be compatible with the observed type shares. At these parameter values, it is possible to construct a network that would deliver the observed type shares and satisfy necessary conditions for pairwise stability. Second, these conditions are also shown to be sufficient for the existence of a pairwise stable network with such type shares in the two important classes of models in Examples 1 and 2 above (subject to a restriction on the parameter space in the case of Example 2). This result then allows us to learn exactly what sets of preference parameters would be consistent with the equilibrium types shares observed in the data. So, for these examples, we are able to characterize the identified sets for the parameters: a parameter belongs to the identified set if and only if at this parameter, there exists a stable network that has the same type shares as are observed in the data.

4.1 Necessary Conditions for Pairwise Stability

This section provides our general result on identification. The theorem below takes as given the proportions of individuals of each network type. It provides necessary conditions for this (observable) distribution of network types to correspond to a pairwise stable network that could be generated under a given preference structure. A given parameterization yields a distribution of preference classes, which enters into the conditions below. The intuition behind these conditions is discussed after the statement of 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). *There are no mutually beneficial links to add between individuals who are distant from each other in the network; i.e., $d(i, j; G) > 2D$. 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,*

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

To denote the vector of type shares, let $\pi \equiv (\pi_t)_{t \in \mathcal{T}}$ be such that $0 \leq \pi_t \leq 1$ for any t and $\sum_{t \in \mathcal{T}} \pi_t = 1$. The element π_t is the proportion of individuals in the network who are of network type t (i.e., the type share for type t). This vector is derived from the observable features of the network: the global adjacency mapping G and the vector of characteristics \mathbf{X} . For a given vector π , we can state the following result.

Theorem 1. *Given a distribution of preference classes in the population, 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 $t \in \mathcal{T}$.*

In the theorem the probabilities of the preference classes are taken as given, as they are determined from the structural preference parameters and distribution of preference shocks. The proof appears further below. First we provide some discussion of the two conditions, which will shed more light on the link between pairwise stability and the observable shares of network types in these classes of models.

Discussion:

Condition 1 is related to expression (1) in the definition of pairwise stability, which pertains to existing links. The purpose of Condition 1 is to require that if there is a link between two individuals, then both are better off being linked than not. In particular, consider any link ij in the network. For this link to be pairwise stable, the following must hold:

$$u(A_i, v_i; \epsilon_i) \geq u(A_{i,-1}, v_i; \epsilon_i) \quad \text{and} \quad u(A_j, v_j; \epsilon_j) \geq u(A_{j,-\mathbf{k}}, v_j; \epsilon_j),$$

where the (local) adjacency matrices $A_{i,-1}$ and $A_{j,-\mathbf{k}}$ are obtained by the deletion of the link between i and j , which appears in row 1 of A_i and row \mathbf{k} of A_j . Condition 1 treats individuals i and j separately, but this of course implies that the above inequalities hold for the pair. If Condition 1 fails and $\alpha_H(t) > 0$ for some $t \notin H$, then there is a positive mass of individuals who would like to drop one of their links, given the definition of H . In this case, the corresponding network cannot be pairwise stable. The practical impact of this condition is fairly obvious, as we illustrate below.

Example (3 cont'd). *In the context of Example 3, Condition 1 allows only 16 out of the 48 allocation parameters to be nonzero. For blacks the 8 potentially nonzero parameters are $\alpha_{H_1}(B, 0), \alpha_{H_2}(B, 0), \alpha_{H_2}(B, B), \alpha_{H_3}(B, 0), \alpha_{H_3}(B, W), \alpha_{H_4}(B, 0), \alpha_{H_4}(B, B), \alpha_{H_4}(B, W)$, where the preference classes H_1 to H_4 are as defined previously. For a vector of preference parameters to be in the identified set, it must produce preference class probabilities (i.e., $P_{H|x}$) that can yield the observed distribution of network types using only these allocation parameters.*

Condition 2 is related to expression (2) in the definition of pairwise stability, which pertains to nonexisting links. This condition establishes that there would be no further mutually beneficial links to add *between individuals who are greater than $2D$ from each other in the network*. Later, in Section 4.2, we show that for the empirically relevant classes of models in Examples 1 and 2, Condition 2 is also necessary and sufficient for nonexisting links between individuals who are $2D$ or less from each other (for a subset of the parameter

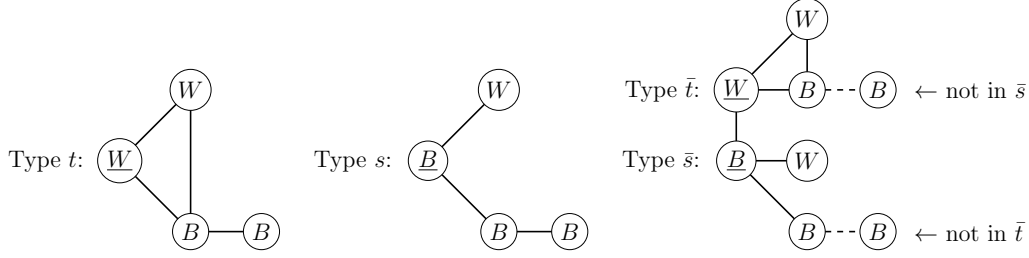


Figure 2: Example of a link added between individuals who are initially distant from each other. *Note:* Dashed lines indicate connections to nodes that appear in only one of the new types, because they are beyond $2D$ from the ego of the other type.

space in the case of Example 2, see Assumption 3).¹⁴ Consequently, Conditions 1 and 2 can guarantee pairwise stability in these models, and hence the approach would yield the sharp identified set. This holds without restriction in Example 1. In Example 2, this holds for parameters satisfying Assumption 3, and so *in that region of the parameter space* the conditions would yield the sharp set.

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 example, consider the types t and s illustrated in Figure 2, which have $D = 2$ as in Example 2. If a link were added between two individuals of these types, who were initially unconnected in the network, they would be transformed to the types \bar{t} and \bar{s} as illustrated. The same transformation would occur if these individuals were initially connected at any distance greater than $2D$. 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, then, because the resulting types do not depend on the exact distance between the two individuals (so long as it is greater than $2D$), the verification of Condition 2 does not require information on the global network G .

Nonexisting links between individuals who are $2D$ or less from each other are not considered in Condition 2, because different transformations could occur in that case. For example,

¹⁴In more general classes of models it is possible to develop a *necessary* condition for nonexisting links between individuals who are within $2D$ from each other, but we have not found a tractable version of such a condition that would also *guarantee* pairwise stability in generality.

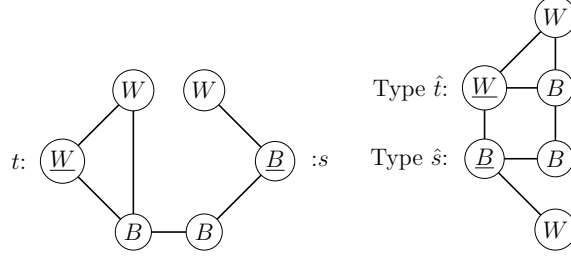


Figure 3: Example of a link added between individuals who are already close in the network

if two individuals of the same types t and s from Figure 2 were initially connected at distance 3, as in Figure 3, 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} . In particular, types \bar{t} and \bar{s} each have three friends of friends while types \hat{t} and \hat{s} each have two, so under the specification in Example 2 the utility of the latter types would be lower by ν . Because Condition 2 focuses on potential links between distant individuals, it obviously carries less information about parameters that moderate incentives to form connections to people who are already nearby, such as the parameter ω in Example 2.

Next, note that with a positive measure of individuals of type s in the network, then there are infinitely many individuals of type s who are beyond $2D$ from any one individual of type t . 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.

Example (3 cont'd). *To apply Condition 2 in our simple model of best-friendships, notice first that only isolated individuals have the capacity to add links. Hence the pairs of types that are relevant for this condition are $(B, 0)$ and $(B, 0)$, $(B, 0)$ and $(W, 0)$, and $(W, 0)$ and $(W, 0)$. Let type $t = (B, 0)$ and $s = (W, 0)$. A link between t and t yields type $\bar{t} = (B, B)$ for both individuals. Hence in the equation for the pair (t, t) in Condition 2, the two measures*

being considered are identical, and so there must be zero measure of isolated blacks who would prefer to add a link to another black. The same follows for isolated whites using the pair (s, s) . This restricts the allocation parameters $\alpha_{H_2}(B, 0)$ and $\alpha_{H_4}(B, 0)$ to be zero, and same for the analogous parameters for whites. For the pair of types (t, s) , a link between them would transform t to (B, W) and s to (W, B) . Hence it must be that either $\alpha_{H_3}(B, W)$ and $\alpha_{H_4}(B, W)$ are zero, or the analogous parameters for whites are zero. This means that there must be either no isolated blacks who would prefer to have white best friends, or no isolated whites who would prefer to have black best friends.

In the proof below we argue that if, under some given distribution of preference classes, there exists a pairwise stable network with the proportions of network types π , then allocation parameters can be found that yield π from this distribution of preference classes and satisfy Conditions 1 and 2.

Proof. Given a 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 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 (1) in Definition 1. Hence all individuals must be of network types that are within their preference classes, because they do not prefer to drop any existing link. Hence $\alpha_H(t) > 0$ only if $t \in H$, and so Condition 1 is satisfied.

If the network is pairwise stable, then all nonexisting links satisfy expression (2) in Definition 1. Partition all nonexisting links (i.e., all pairs of individuals (i, j) such that $G(i, j) = 0$) by the pair of types of the two individuals (i.e., (t, s) where $t_i = t$ and $t_j = s$). Consider an arbitrary such pair of types (t, s) where both types have fewer than L links (the maximum). Let (\bar{t}, \bar{s}) be the pair of types that 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 existing network G . Suppose a positive measure of individuals of type t would prefer to add links to some individuals of type s who are greater than $2D$ away from them. If so, these individuals of type t would prefer to add links to any individuals of type s who are greater than $2D$ away (because adding a link to any such individual of type s would yield the same marginal payoff to a given individual of type t). In terms of Condition 2, this corresponds

to there being a positive measure of individuals of type t who also have \bar{t} in their preference classes (i.e., $1_{\bar{t} \in \tilde{H}}$). Because the network is pairwise stable, none of the individuals of type s would prefer to add links to the individuals of type t who want to link with them. Given that there is a positive measure of individuals of type t who want to add such links, for every individual of type s there must be an individual of type t who is greater than $2D$ away and wants to link with her. Furthermore, because the marginal payoff to an individual of type s from adding link to an individual of type t is the same regardless of which individual of type t is used (so long as that individual is beyond $2D$), it must be that no individuals of type s would prefer to add links to any individuals of type t who are greater than $2D$ away from them. Hence the measure of individuals who are type s but who also have \bar{s} in their preference classes must be 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. \square

We provide two remarks, on the generality of the theorem and on why the conditions are not also sufficient to guarantee pairwise stability (in general).

Remark 3. *The theorem applies to any model that satisfies Assumptions 1 and 2. Hence the approach developed here can be used in any application where it is reasonable to assume a finite D and L , have finite cardinality of \mathcal{X} , and have preference shocks pertain to the characteristics of alters rather than their identities. Furthermore, as suggested earlier, one can extend the preference shocks so that there is one for every alter in the network type, not just the direct connections. The number of preference classes may expand, but none of the arguments in the above proof would change.*

Additionally, this result can be used with models where nonexistence is possible. If a particular parameterization cannot generate a pairwise stable network (either with the observed type shares, or with any type shares), 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 distribution of types 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.

Remark 4. *Conditions 1 and 2 do not guarantee the existence of a pairwise stable network, in general, because they do not consider nonexisting links between individuals who are already close to each other in the network. If the payoff from adding a direct link to someone who is*

already close to you (within $2D$) is potentially different than the payoff from adding a link to someone who is more distant (beyond $2D$), there may be further restrictions on the set of preference parameters in order for them to be consistent with the data. This would make the identified set smaller than what is obtained with Conditions 1 and 2 alone.

4.2 Necessary and Sufficient Conditions

Here we prove that Conditions 1 and 2 are sufficient (as well as necessary) for the existence of a pairwise stable network with the observed proportions of network types, under the model structures provided in Examples 1 and 2. Whereas we show that this is the case for a subset of the parameter space in Example 2 (see Assumption 3 below), sufficiency holds for any parameter value in Example 1. Hence, when using the conditions to verify whether the data is compatible with equilibrium behavior at a particular parameter value, our approach yields the identified set in Example 1 and (for a subset of the parameter space) in Example 2.

To prove sufficiency in these two classes of models we use a contrapositive argument. Accordingly the premise for the argument is that under the distribution of preference classes given by a particular parameterization of the model, there is no pairwise stable network with the observed distribution of network types. We then show how this implies that, for any network where the distribution of network types matches the observed distribution, one of our conditions must be violated. In what follows, we first give this argument in general terms, and then provide further requirements and details that are specific to each of the two classes of models.

To start, fix a parameter vector and, hence, a distribution of preference classes, as well as the observed distribution of network types. Under these preferences, any network with this distribution of types is unstable. Therefore, for any such network there must be a positive measure of pairs of individuals for whom the presence or absence of a link between them is unstable. To translate this into our aggregate 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 in this class are some network type that is not in H . Therefore $\alpha_H(t) > 0$ for some $t \notin H$ and Condition 1 would be violated.

For nonexistent links ($G(i, j) = 0$), we first consider 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 a positive measure of individuals of these types would prefer to add links with each other. A link between two individuals of types t and s who are distant ($> 2D$) 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. By definition, if an individual of type t prefers to add a link to an individual of type s and thereby become type \bar{t} , then \bar{t} is in that individual's preference class. Hence $P_{\tilde{H}|v_1(t)}\alpha_{\tilde{H}}(t)1_{\bar{t} \in \tilde{H}}$ is strictly positive for at least one preference class \tilde{H} . The same holds for type s : for at least one preference class \tilde{H} , $P_{\tilde{H}|v_1(s)}\alpha_{\tilde{H}}(s)1_{\bar{s} \in \tilde{H}}$ is strictly positive. Therefore the product of the measures given by these expressions is strictly positive, which violates Condition 2.

Last, we consider nonexistent links between individuals who are $2D$ or less from each other in the network. Here we show that if there is a positive measure of unlinked pairs of individuals who are $2D$ or less from each other and who would prefer to be directly linked, then there must also be a positive measure of unlinked pairs of individuals who are greater than $2D$ from each other and who would prefer to be linked. Consequently Condition 2 would be violated based on the above argument. Establishing this involves arguments and requirements that are specific to each of the example classes of models.

Sufficiency in Example 1

Only direct connections affect utility in this class of models (i.e., $D = 1$), and the connections among one's friends do not matter. Hence adding a link to an individual of some type s who is distant in the network ($d(i, j; G) > 2D$; i.e., $d > 2$) yields the same marginal payoff as adding a link to someone of the same type who is nearby but not directly connected ($d(i, j; G) = 2$). So if an individual of type t would prefer to add a link to an individual of type s who is at distance 2, which would transform her to some type \hat{t} , she would also prefer to add a link to anyone else of type s who is greater than distance 2 away. This would transform her to some other type \bar{t} , which appears in the equation for types t and s in Condition 2. Thus if a positive measure of individuals of type t would prefer to add links to nearby alters of type s , there is also a positive measure who would prefer to add links to

distant individuals of type s . The same holds for individuals of type s : if a positive measure of them would prefer to add links to nearby alters of type t , there is also a positive measure who would prefer to add links to distant individuals of type t . Hence the equation for the tuple $(t, s), (\bar{t}, \bar{s})$ in Condition 2 would be violated.

Thus, if the network is unstable, one of our conditions must be violated. Therefore Conditions 1 and 2 are sufficient to guarantee pairwise stability in this class of models.

Sufficiency in Example 2

For this class of models, we require the following assumption.

Assumption 3. *The value of mutual friends is not excessively large relative to the value of friends of friends: $\omega \leq \frac{L}{L-1}\nu$. Also the value of friends of friends is non-negative: $\nu \geq 0$.*

Under this assumption we can show that having a preference to add a link to a nearby alter (i.e., $d \leq 2D = 4$) implies a preference to add a link to someone of the same type at any distance greater than $2D$.

Consider the marginal utility to individual i from adding a link to individual k of some type s . Under the preference specification in Example 2, if the distance between i and k is greater than $2D$, the marginal utility to i is $f(x_i, x_k) + \epsilon_{il}(x_k) + |N(k)|\nu$. That is because there is no intersection between i 's existing friends of friends (i.e., $\cup_{l=1, j(l) \neq \emptyset}^L N(j(l))$) and k 's friends ($N(k)$). The marginal utility is the same if $d(i, k; G) = 2D = 4$, because the distance between i 's direct friends and k would be at least 3, so again there is no intersection between i 's existing friends of friends and k 's friends. If instead $d(i, k; G) = 3$ then it must be that $N(k) \cap \left(\cup_{l=1, j(l) \neq \emptyset}^L N(j(l))\right) \neq \emptyset$, meaning some of k 's friends overlap with i 's existing friends of friends. An example of this was shown in Figure 3. In this case the marginal utility of adding a link to individual k would be less than the marginal utility of adding a link to someone of type s who is more distant, because i gains fewer new friends of friends.

Finally, suppose that $d(i, k; G) = 2$. In this case the marginal utility of adding a direct 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.¹⁵ 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

¹⁵There could be multiple paths of length 2 from i to k , so there could be multiple new mutual friendships.

be a friend of friend. So there would be $\sum_{l=1, j(l) \neq \emptyset}^L G(j(l), k) + 1$ fewer friends of friends in total, in the resulting type. The difference in the marginal utility of adding a link to this individual of type s who is at distance of 2, compared with adding a link to someone of the same type but who is beyond $2D$, would be

$$\sum_{l=1, j(l) \neq \emptyset}^{L-1} G(j(l), k)(\omega - \nu) - \nu.$$

The maximum possible number of new mutual friendships is $L - 1$, so Assumption 3 guarantees that this difference is weakly negative: $(L - 1)(\omega - \nu) - \nu \leq 0$. Therefore if individual i prefers to add a link to an individual of type s who is at distance 2, she would also prefer to add a link to any other individual of the same type at any distance $> 2D$.

So, as in Example 1, if a positive measure of individuals of some type t would prefer to add links to nearby alters of some type s , they would also prefer to add links to any distant individuals of type s . The same holds for the individuals of type s . Therefore if there is a positive measure of individuals of some pair of types t and s who would prefer to add links with their nearby alters of the corresponding types, then there is also a positive measure of individuals of types t and s who would prefer to add links with more distant individuals of the same types. This would violate Condition 2. Combining this with the previous arguments for existing links and for nonexisting links between individuals greater than $2D$ from each other, we can conclude that Conditions 1 and 2 are sufficient to guarantee pairwise stability in this class of models, under Assumption 3.

Thus we have established the following result.

Theorem 2. *For the classes of models in Examples 1 and 2 (satisfying Assumption 3 in the case of Example 2), given a distribution of preference classes in the population, there exists a pairwise stable network where the proportion of agents of type t is equal to π_t for each $t \in \mathcal{T}$ 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 $t \in \mathcal{T}$.*

5 Implementation

In this section we describe how to implement our approach to identification with network types. We establish that the conditions previously presented for an observed network to correspond to a pairwise stable network with pre-specified structural parameters can be verified using a quadratic program. A brief discussion of statistical inference is given in the Appendix.

5.1 Formulation as Quadratic Programming Problem

Condition 2 provides a quadratic function of the allocation parameters, which is minimized at zero, subject to additional linear constraints, if the data rationalize a pairwise stable network.¹⁶ The constraints require that the predicted proportions of types must match the observed proportions, and that individuals are only allocated to types in their preference classes (Condition 1). The latter can be accomplished simply by omitting the allocation parameters for types not in a given preference class.

To assemble the quadratic program, we first define a matrix Q that can be used to express the quadratic function for Condition 2. This is a square matrix that has one row and column for each allocation parameter (e.g., $\alpha_H(t)$) whose type (t) appears in that preference class (H). Thus there are only rows and columns for allocation parameters that satisfy Condition 1. Briefly, then, the elements of Q will identify which pairs of these allocation parameters could result in the network having unlinked individuals who want to add links with each other. Note that each parameter $\alpha_H(t)$ corresponds to individuals with preferences in a particular class H who have been assigned to a particular type t . So a pair of allocation parameters, say $\alpha_H(t)$ and $\alpha_G(s)$, corresponds to two sets of individuals with preferences in classes H and G and who are types t and s , respectively. It may be the case that it would be both feasible and mutually beneficial for individuals of these types and with these preferences to add links with each other. In such cases, Condition 2 requires that the measure of individuals allocated to one or the other of these combinations must be zero.

To construct Q , we start with a square matrix S of the same dimension. The row for a given parameter $\alpha_H(t)$ in S indicates which allocation parameters (i.e., which columns) correspond to any types with whom individuals of type t with preferences in class H would want to add links (if those alters are greater than $2D$ away). These would be any types s

¹⁶A similar approach to “solving” the identified set is Honoré and Tamer (2006)’s use of a linear program in a nonlinear panel data model.

such that adding a link with an individual of type s (who is greater than $2D$ away) would transform an individual of type t to some type \bar{t} , where $\bar{t} \in H$. Because $\bar{t} \in H$, an individual of type t with these preferences would want to add this link. The row for $\alpha_H(t)$ accordingly has entries equal to 1 in the columns for all parameters $\alpha_G(s)$ where s is one such type, and 0 otherwise.

The matrix Q equals the Hadamard (i.e., entrywise) product of S with its transpose S^\top : $S \circ S^\top$. Think of each element of Q as involving two combinations of (preference class \times type), say (H, t) and (G, s) . The row in matrix S referring to preference class H and type t indicates whether individuals in that preference class and of that type would prefer to add links to individuals of type s . The matrix S^\top tells us whether those individuals of type s who are in preference class G would also prefer to add links to individuals of type t . Thus this element of $S \circ S^\top$ tells us whether additional links would be mutually preferred between individuals from the two corresponding elements of the allocation, (H, t) and (G, s) .

Condition 2 requires that the measure of individuals from at least one of these two elements of the allocation be zero. Otherwise there would be mutually preferred links to add between individuals who are greater than $2D$ from each other. This requirement is equivalent to $\alpha_H(t)\alpha_G(s) = 0$. Then, because $\alpha_H(t) \geq 0$ and $\alpha_G(s) \geq 0$ for all (H, t) and (G, s) , $\alpha^\top Q \alpha = 0$ expresses this requirement for all such pairs of allocation parameters. (To match Condition 2 exactly, we would include the measures of individuals in the relevant preference classes, e.g., $\mu_{v_1(t)} P_{H|v_1(t)}$, in S and Q . As we point out below, though, this is equivalent to Condition 2 because $\mu_{v_1(t)} > 0$ and $P_{H|v_1(t)} > 0$.)

Example (3 cont'd). *To illustrate the objective matrix Q , consider our simple example with direct links only ($D = 1, L = 1$) and two possible characteristics ($\mathcal{X} = \{B, W\}$). As noted previously, in this example types can be described using just the vectors of characteristics v . The preference classes can then 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 $H_8 = \{(W, 0), (W, W), (W, B)\}$. If we omit the allocation parameters for any types that are not in a preference class, there are 16 remaining parameters: $\alpha_1(B, 0)$; $\alpha_2(B, 0)$, $\alpha_2(B, B)$; $\alpha_3(B, 0)$, $\alpha_3(B, W)$; $\alpha_4(B, 0)$, $\alpha_4(B, B)$, $\alpha_4(B, W)$; $\alpha_5(W, 0)$; $\alpha_6(W, 0)$, $\alpha_6(W, W)$; $\alpha_7(W, 0)$, $\alpha_7(W, B)$; $\alpha_8(W, 0)$, $\alpha_8(W, W)$ and $\alpha_8(W, B)$. The subscripts of the parameters correspond to the subscripts of the preference classes. Arrange these parameters into a vector $\alpha = (\alpha_1(B, 0), \alpha_2(B, 0), \alpha_2(B, B), \dots, \alpha_8(W, B))$.*

The matrix Q is shown in Figure 4. We note that both S and Q are sparse binary

(row)	(parameter)	
1	$\alpha_1(B, 0)$	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 0
3	$\alpha_2(B, B)$	0 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 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 0
6	$\alpha_4(B, 0)$	0 1 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0
7	$\alpha_4(B, B)$	0 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 0
9	$\alpha_5(W, 0)$	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 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 0
12	$\alpha_7(W, 0)$	0 0 0 1 0 1 0 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 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 0
16	$\alpha_8(W, B)$	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

Figure 4: Example of Q

matrices, which greatly reduces the memory requirements for computation. Finally we can see how $\alpha^\top Q \alpha = 0$ expresses Condition 2. In this example it yields

$$\begin{aligned} \alpha^\top Q \alpha = & \alpha_2(B, 0)^2 + 2\alpha_2(B, 0)\alpha_4(B, 0) + 2\alpha_3(B, 0)\alpha_7(W, 0) + 2\alpha_3(B, 0)\alpha_8(W, 0) + \alpha_4(B, 0)^2 \\ & + 2\alpha_4(B, 0)\alpha_7(W, 0) + 2\alpha_4(B, 0)\alpha_8(W, 0) + \alpha_6(W, 0)^2 + 2\alpha_6(W, 0)\alpha_8(W, 0) + \alpha_8(W, 0)^2. \end{aligned}$$

(As noted earlier, this does not include the measures of individuals in the relevant preference classes, i.e., $\mu_x P_{H|x}$, but these are strictly positive.) Not all allocation parameters appear in the expression above (e.g., $\alpha_1(B, 0)$ or $\alpha_8(W, B)$), so $\alpha^\top Q \alpha = 0$ does not require that the vector α is all zeros. Typically the matrix Q will not be positive definite, so its null space will include more than a vector of zeros.

Once Q is assembled as illustrated above, we can show the following.

Theorem 3. *Given a structural parameter vector θ yielding a distribution of preference*

classes $P(\cdot)$, a network with type shares $\{\pi_x(t)\}$ satisfies Conditions 1 and 2 if and only if

$$\min_{\{\alpha_H(t): t \in H\}} \alpha^\top Q \alpha$$

subject to:

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

$$\alpha_H(t) \geq 0, \forall t, H$$

$$\sum_H P_{H|x} \alpha_H(t) = \pi_x(t), \forall t$$

is equal to zero.

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

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

The first two sets of constraints in the program simply require that allocations from a given preference class add up to one and are non-negative. The third set of constraints is to match the observed proportions of network types ($\pi_x(t)$). 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. \square

Remark 5. One could alternatively rely on a quadratic form that exactly reproduces Condition 2, by including the probabilities $P(\cdot)$ and measures $\mu_{(\cdot)}$ in the elements of Q . The configuration above, however, does not require one to recompute Q at each putative parameter vector $P(\cdot)$, and as noted above the matrix is a sparse, binary matrix which saves memory in some programs such as Matlab.

The theorem above assumes the population type shares are known. In order to accommodate data from finite networks, we modify the QP problem to allow for error around the observed shares.¹⁷ To do this we introduce two slack variables for each type share, one for a positive difference ($\beta^+(t)$) and one for a negative difference ($\beta^-(t)$), so that the third set of constraints becomes¹⁸

$$\sum_H P_{H|x} \alpha_H(t) + \beta^+(t) - \beta^-(t) = \pi_x(t), \forall t.$$

The slacks are additional variables in the modified QP problem, but they do not enter into the objective function (it remains the same). Instead the magnitudes of the slacks are constrained based on functions of the sample size n , denoted $\delta^+(n)$ and $\delta^-(n)$, such that $0 \leq \beta^+(t) \leq \delta^+(n)$ and $0 \leq \beta^-(t) \leq \delta^-(n)$, $\forall t$. These constraints define hard “bands” around the observed type shares.¹⁹ Accordingly, the modified QP problem verifies whether a structural parameter vector can yield a prediction within these bands around the observed shares while satisfying Conditions 1 and 2. (Appendix B.2.8 provides further details in a particular example.)

Finally, we note that the objective function $\alpha^\top Q \alpha$ may be nonconvex (because while the matrix Q is symmetric, it may be indefinite as is the case with the example in Figure 4), which poses a problem for many standard QP solvers. A more general constrained nonlinear optimization routine can be used instead. Importantly, the fact that the optimal value is known (i.e., $\alpha^\top Q \alpha = 0$) makes it trivial to verify that a global rather than local optimum has been reached. In the simulation examples in the next section, we find that the QP problem solves easily using an active set algorithm in the program KNITRO.

6 Simulations

We now present two simulation exercises based on the example classes of models discussed throughout the paper. The main purpose is to illustrate the performance of our approach,

¹⁷This could be viewed as an adjustment to enable the continuum model to be used as an approximation for finite networks.

¹⁸The use of two slack variables rather than one is a standard approach to minimize the sum of absolute deviations via linear programming. We use a preliminary linear programming problem to find starting values for the QP (see Appendix B.2.8).

¹⁹If a sampling distribution for the observed vector of type shares were available, 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 A.

in terms of the parameter sets that are recovered and the computational burden that is involved. Additionally, the procedures described here and in the Appendix provide some guidance on further aspects of implementation, such as how to generate the sets of network types and preference classes, and how to construct the matrix Q .

6.1 Simple Model with One Link (Example 3)

The first exercise uses the simple model introduced in Section 2.3 and later referred to as Example 3. To review the features of this model, individuals have at most one link and the predetermined characteristics are B (“black”) and W (“white”); thus $D = 1$, $L = 1$, and $|\mathcal{X}| = 2$. Types can be fully characterized with the vector $v = (x, y)$, where x is the characteristic of the ego and y is the characteristic of the alter (or 0 if the ego is isolated). Given the utility specification in (3), the probabilities p_{xy} defined in (4) can serve as the primitives for the distribution of preferences (e.g., $p_{BW} = 1 - F_\epsilon(-f_{BW})$). The probabilities of preference classes then have simple expressions, such as $P_{H|B} = p_{BB}(1 - p_{BW})$ for $H = \{(B, 0), (BB)\}$.

In this simple model, it is relatively easy to find *all* the equilibrium vectors of type shares for a given vector of structural parameters (see Appendix B.1). This is useful for demonstrating the full range of possible equilibria under a single parameterization. Figure 5 presents the set of equilibria for $p_{BB} = 0.4$, $p_{BW} = 0.2$, $p_{WB} = 0.15$, $p_{WW} = 0.5$, with population sizes $\mu_B = 1.0$, $\mu_W = 1.2$. There are six types in total ($(B, 0)$, (B, B) , (B, W) , $(W, 0)$, (W, B) , (W, W)), but the shares of two are redundant (e.g., $\pi_{(B,0)} = 1 - \pi_{(B,B)} - \pi_{(B,W)}$). The figure shows different margins of the remaining four-dimensional set. The top-left plot has the shares of blacks linked to blacks ($\pi_{(B,B)}$) and blacks linked to whites ($\pi_{(B,W)}$). High values of both of these shares are not possible because individuals can have at most one link. The top-right plot has the analogous shares for whites ($\pi_{(W,W)}$ and $\pi_{(W,B)}$). In the bottom-left plot we see the linear relationship that must hold between the shares of blacks linked to whites and whites linked to blacks, because links are one-to-one; i.e., $\mu_B \pi_{(B,W)} = \mu_W \pi_{(W,B)}$. Figure 5 also shows the type shares from one finite network (with 500 blacks and 600 whites), generated with a microsimulation procedure we use mainly for the second simulation exercise (see the next section).

We then apply the quadratic programming approach developed above, using the type shares from this finite network (details of the procedures are described with the second simulation exercise). Figure 6 shows the resulting identified set. The values of the cross-

race preference probabilities (p_{BW} and p_{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 probabilities (p_{BB} and p_{WW}) are fairly small and informative: 0.38–0.51 for blacks and 0.46–0.57 for whites. However the identified set would not provide conclusive evidence on the ranking of these parameters (i.e., the fact that $p_{WW} > p_{BB}$).

We also use a simplified version of the quadratic program, which can be derived in this particular case (see Appendix B.1), to confirm the identified set obtained above and to find the identified sets that would be obtained from other vectors of type shares. The simplified QP is trivial to solve, so we can evaluate a large number of parameter vectors almost instantaneously. This makes it easy to explore how the results would change with different observations (i.e., different type shares). For the type shares from the finite network, the identified set found using the simplified QP matches the identified set found above (see Figure A1). We then select four vectors of type shares at random from the set of all equilibria in Figure 5, and qualitatively similar identified sets are recovered from each of these observations as well (see Figure A2 for the 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. That is because the parameters must be able to predict all four vectors of type shares as equilibrium outcomes. (Recall that each network is large, and we assume we can consistently estimate the type shares in each one.) As Figure 7 shows, the identified set from all four networks together is quite precise. Only one parameter vector in the grid we evaluate is accepted ($p_{BB} = 0.40$, $p_{BW} = 0.20$, $p_{WB} = 0.16$, $p_{WW} = 0.50$), which means the identified set spans less than 0.04 in each dimension (the grid uses intervals of 0.02). This illustrates the possibility for substantial identifying power from the observation of multiple large networks.

6.2 Model with Indirect Connections (Example 2)

Next we consider a specification of the more elaborate class of models in Example 2, where $D = 2$. For this exercise the maximum number of links is set to $L = 3$. To help motivate this, 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 in the previous exercise: black or white race, $\mathcal{X} = \{B, W\}$.

The ratio of the measures of blacks to whites (μ_B/μ_W) is set at 1/4.

In terms of A and v (rather than G and \mathbf{X}), the utility specification can be written as follows:

$$\begin{aligned}
u(A, v; \epsilon) \equiv & \sum_{l=2}^{L+1} a_{1l} (f_{v_1, v_l} + \epsilon_{l-1}(v_l)) && \text{(direct connections)} \\
& + \nu \sum_{k>L+1} 1 \left\{ \sum_{l=2}^{L+1} a_{1l} a_{lk} > 0 \right\} && \text{(friends of friends)} \\
& + \omega \sum_{l=2}^{L+1} \sum_{k>l}^{L+1} a_{1l} a_{lk} && \text{(mutual friends)}
\end{aligned} \tag{5}$$

(recall that row 1 of A and v corresponds to the ego of the type, rows 2 to $L + 1$ correspond to direct connections, and rows $k > L + 1$ correspond to friends of friends).²⁰ The preference parameters are chosen to generate a degree distribution that is similar to the observed distribution in Add Health, and the values of ν and ω satisfy Assumption 3.²¹ The preference shocks are drawn from a standard normal distribution, and are then ordered within alter characteristic so that $\epsilon_{i1}(B) > \epsilon_{i2}(B) > \epsilon_{i3}(B)$ and $\epsilon_{i1}(W) > \epsilon_{i2}(W) > \epsilon_{i3}(W)$. This ordering ensures that if an individual would be willing to have l friends of race x (with no mutual connections among them), she would also be willing to have $l - 1$ friends of that race. We think this is more realistic than the alternative, where for example someone who is willing to have two friends of some race might also prefer being alone over having one friend of that race. Additionally, this ordering of the shocks reduces the number of preference classes.

To generate the data we simulate a number of finite networks with $n = 500$ individuals (100 blacks and 400 whites), using a procedure described in Appendix B.2.1. Only one network is needed to produce the observed type shares. It is useful, however, to have a sense of the variation that can arise under this specification for a fixed vector of preference parameters but with different realizations of the preference shocks and different equilibria. Figure 8 plots the shares of certain types or combinations of types appearing in these simulated networks. One network, selected at random, serves as the observation we then use to compute the identified set. The type shares in this network (indicated with triangles in Figure 8) are near

²⁰The $1\{\cdot\}$ in the second line of (5) indicates whether a path of length two exists from the ego to the indirect alter k .

²¹The 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.

the center of the distribution of shares shown in the figure. Additionally, we see that the shares of types with any mutual friends are in fact zero in most simulated networks. 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 $f_{xy} + \epsilon_l(y)$).²²

The specific procedures used to formulate and solve the QP are described in Appendix B.2. Here we note two key points. First, 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). There are 68 observed types in the selected network and an additional 72 unobserved but adjacent types, while the preference specification yields a total of 356 possible types. Second, the QP used in this exercise includes an additional constraint on the sums of the slack variables ($\sum_t \beta^+(t)$ and $\sum_t \beta^-(t)$). Without this, the total absolute error between the model prediction and the observed type shares would be equal to the size of the band around each individual type share (i.e., $\delta^+(n)$ or $\delta^-(n)$) multiplied by the number of types in the QP (in this case, $68 + 72 = 140$). Given the large number of types, even a small band size like 0.01 would 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.

We use a Markov Chain Monte Carlo (MCMC) algorithm to generate the identified set, which is represented in Figures 9 and 10. Figure 9 plots identified values of 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 for 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

²²Given $\omega = \nu$, 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.

the previous example.²³ Figure 10 plots the identified values of ν (friends of friends) against the parameters for direct connections. At the lower boundary there is an inverse relationship between ν and each of the f_{xy} , likely because in the observed network most links also provide one or more indirect connections, so the relevant marginal value in most cases is $f_{xy} + \nu$. The identified range of ν itself 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). 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 follows from Assumption 3: $\omega \leq \frac{3}{2}\nu$.

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. Furthermore, the MCMC 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.

7 Conclusion

This paper provided an approach to identification of preferences in models of network formation using data on one large network or a sequence of large networks. Given a preference structure, and using pairwise stability as the basis for strategic network formation, we derive necessary conditions that the true parameters must satisfy under a set of reasonable assumptions on model primitives. These necessary conditions involve allocating individuals to “network types” in a way that matches the observed distribution of these types, and is compatible with pairwise stability. We then show that our conditions are sufficient (and necessary) in two classes of models that have been studied in the literature. In these cases, our conditions characterize the identified set. We then provide a quadratic programming algorithm that we use to construct this set and illustrate how this is done using Monte Carlo simulations.

²³In 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 novelty of our approach stems from using the economic model under pairwise stability as a vehicle for effective dimension reduction. In particular, restricting depth and number of links via assumptions on D and L , along with the x 's, defines a finite set of network types that are relevant for the preference structure. These types are observed in the data and present the crucial link that we use to learn about preference parameters. Our simulation exercises indicate that the identified sets are indeed informative, and that the approach is feasible for empirically realistic models.

In this paper, we focused our work on setting up a model of network formation under pairwise stability and studied how one might recover its identified features under our maintained assumptions. We have also provided a computational algorithm that one can use to learn about these parameters of interest. An obvious unfinished question is effective statistical inference given the identification analysis we have done. In the Appendix, we show that given a limit assumption on the data, it is simple to translate statistical uncertainty about type shares that are observed in the data to uncertainty about preference parameters. However, providing sufficient conditions for these limit theorems to hold is beyond the scope of this current paper, and it is something that we leave for future work.

References

- BACKSTROM, L., P. BOLDI, M. ROSA, J. UGANDER, AND S. VIGNA (2012): “Four Degrees of Separation,” <http://arxiv.org/pdf/1111.4570.pdf>.
- 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 (2013): “My Friend Far Far Away: Asymptotic Properties of Pairwise Stable Networks,” Working Paper.
- 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): “Tractable and Consistent Random Graph Models,” Working Paper 20276, NBER.
- CHIAPPORI, P.-A., A. GALICHON, AND B. SALANIÉ (2012): “The Roommate Problem Is More Stable Than You Think,” 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.
- CURRARINI, S., M. JACKSON, AND P. PIN (2009): “An Economic Model of Friendship: Homophily, Minorities and Segregation,” *Econometrica*, 77(4), 1003–1045.
- DZEMSKI, A. (2014): “An Empirical Model of Dyadic Link Formation in a Network with Unobserved Heterogeneity,” University of Mannheim 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.

- GRAHAM, B. (2014): “An Empirical Model of Network Formation: Detecting Homophily when Agents are Heterogeneous,” University of California at Berkeley Working Paper.
- HASTIE, T., R. TIBSHIRANI, AND J. FRIEDMAN (2001): *The Elements of Statistical Learning*. Springer.
- HONORÉ, B. E., AND E. TAMER (2006): “Bounds on Parameters in Panel Dynamic Discrete Choice Models,” *Econometrica*, 74(3), 611–629.
- 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.
- KALAI, E. (2004): “Robust Large 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.
- KOLACZYK, E. (2009): *Statistical Analysis of Network Data*. Springer-Verlag.
- 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. (2013): “A Structural Model of Segregation in Social Networks,” Johns Hopkins Working Paper.
- MENZEL, K. (2014a): “Inference for Games with Many Players,” NYU Working Paper.
- (2014b): “Large Matching Markets as Two-Sided Demand Systems,” NYU Working Paper.
- MIYAUCHI, Y. (2014): “Structural Estimation of a Pairwise Stable Network with Nonnegative Externality,” MIT Working Paper.
- POLITIS, D., J. ROMANO, AND M. WOLF (1999): *Subsampling*. Springer.
- SHENG, S. (2014): “A Structural Econometric Analysis of Network Formation Games,” UCLA Working Paper.
- SONG, K. (2014): “Econometric Inference on Large Bayesian Games with Heterogeneous Beliefs,” UBC Working Paper.

A Statistical Inference

As the paper’s main contribution is identification analysis in large networks, we do not focus on the estimation problem. The estimation problem in our setup is fraught with difficulties having to do with interdependencies in links and link formations. For example, we cannot simply assume that the absence of a link between two individuals makes the nodes independent. This absence can simply be a result of a decision by the two individuals not to link. Under assumptions stated below, we nevertheless provide an avenue that one can use for constructing confidence regions.

The parameter of interest in our setup is the vector that characterizes the payoff structure, θ . On the other hand, the data are informative only on the measure of network types, $\Pi^\top \equiv (\pi(t_1), \dots, \pi(t_{|\mathcal{T}|}))$. Since the network types are observed, these can be interpreted as “choice probabilities”. The setup leads itself, heuristically, to a method of simulated moments approach where for a given value of the parameter θ , we can simulate a vector of *predicted* Π ’s as a function of θ , and then minimize an appropriate distance between the predicted Π and the estimated Π . The main insight is that, *knowing* Π , *we can solve for the identified set for* θ . We do not know Π , but can estimate it using the data. Hence, the question becomes one of mapping the statistical uncertainty about Π to uncertainty about the identified set for θ .

Let there be a given vector Π of observed type probabilities. Then, the identified set $\Theta \subset \mathbb{R}^k$ in a given (large) network can be defined as follows (without conditioning on \mathbf{X}):

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

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 5 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 $F(\theta, \Pi) = 0$. Heuristically, to obtain a confidence region for Θ , we can first construct a confidence region for Π and then “map” this set to a confidence region for Θ (which can be a set of sets). This heuristic relies fundamentally on being able to construct a valid confidence region for Π . This is a hard problem that we mostly leave for future research. Whether we are able to construct such a confidence region for Π (or credible set) depends on the kind of data we have. For example, if we observe a large sequence of i.i.d. networks each of which is large, then it is simple to see that one can treat each network as an “observation” that provides a copy of Π and then we can use such a large sequence under standard limit theorems for i.i.d. observations (each observation here would be a finite vector of type probabilities). Unless we assume that the same equilibrium is being “played” in every market, however, we must confront the issue of multiple equilibria (see Remark 7).

A slightly more realistic data design is one in which we observe one large network²⁴ and we can think of this network as a draw from the data generating process. This design is similar to standard time series models where conditions such as stationarity and ergodicity induce the kind of (conditional) exchangeability needed for learning via limit theorems. The next section highlights briefly how this is done given such limit theorems in a large network.

A.1 Inference with a Large Network

Suppose that our data is given by a large matrix of connections among n individuals from which we are able to determine the network type t of each individual i in the sample. (Assume away regressors for simplicity.) This will allow us to have a sample analogue of the measure for each type,

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

for $k = 1, \dots, |\mathcal{T}|$ and where these types are mutually exclusive. Moreover, let $\hat{\Pi}^\top = (\hat{\pi}(t_1), \dots, \hat{\pi}(t_{|\mathcal{T}|}))$ which is the vector of estimated type probabilities. We make the following

²⁴Statistical uncertainty can alternatively be due to partial observability of the network. In general this is an important limitation. Nevertheless, given a known sampling scheme, it is possible to extract information about underlying features of a network and hence the network type distribution (see, e.g., Kolaczyk (2009)). In this case, the fact that we do not need to rely on the observation of the whole network to estimate the proportion of network types is an advantage of our method. In addition, observing a sample of the network using a known sampling scheme allows for easier inference on the true type shares even in the presence of correlations (as in Kolaczyk (2009)).

assumption on the population choice probabilities.

Assumption 4. *Let the choice probabilities be such that for some $\delta > 0$*

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

In addition, the following goodness of fit statistics is such that, as $n \rightarrow \infty$

$$G(\hat{\Pi}, \Pi) = n \sum_{k=1}^{|\mathcal{T}|} \frac{(\hat{\pi}(t_k) - \pi(t_k))^2}{\hat{\pi}(t_k)} \rightarrow_d \chi_{|\mathcal{T}|}^2, \quad (6)$$

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

Remark 6. *Note that the fact that we maintain a χ^2 limiting distribution is not essential. If we instead maintain the much weaker condition that $a_n(\hat{\Pi} - \Pi) \rightarrow^d Z$ where Z is a nondegenerate random variable and a_n is a sequence of non-stochastic positive constants that tends to infinity, then one can still use subsampling based approaches for inference. (This is possible even if a_n is not known. See, e.g., Politis, Romano, and Wolf (1999).) But, here, for clarity we maintain such chi-squared limit. Of course in both cases, the main strength of the assumption, and something that we do not verify here, is that this sequence of random variable does converge to a nondegenerate limit (be it Z or χ^2). Requiring that the probabilities are bounded away from zero is not strong as one can redefine the types in a way that guarantee that the condition holds. Providing lower level conditions that guarantee that (6) holds is beyond the central contribution of the paper, which is focused on building a computationally tractable approach to identification in these models.*

Given the above assumption, to build a (frequentist) confidence region for Θ , one can use a projection method as follows.²⁵ First, we construct a confidence region for the type probabilities. This is a standard inference problem under the above assumption. This can be done for example, by inverting some test statistic for multinomial probabilities such as a χ^2 test. In particular, define

²⁵Alternative approaches exist, but we choose projections for illustration. Projection methods have been used before in both statistics and econometrics and here we only give a snapshot of projections as our focus is mainly on identification.

$$CI_{1-\alpha}(\Pi) = \{\Pi \in S^{|\mathcal{T}|} : G(\hat{\Pi}, \Pi) \leq c_{1-\alpha}(\chi_{|\mathcal{T}|}^2)\} \quad (7)$$

where $S^{|\mathcal{T}|}$ is the unit simplex of size $|\mathcal{T}|$, $\hat{\Pi}(t)$ are the sample analogues of the type probabilities, $c_{1-\alpha}(\chi_{|\mathcal{T}|}^2)$ is the $(1 - \alpha)$ critical value of the $\chi_{|\mathcal{T}|}^2$ distribution, and finally the goodness of fit statistic is

$$G(\Pi_1, \Pi_2) = n \sum_{k=1}^{|\mathcal{T}|} \frac{(\pi_2(t_k) - \pi_1(t_k))^2}{\pi_1(t_k)}.$$

The confidence region in (7) 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:

$$CI_{1-\alpha}(\theta) = \{\Theta(\Pi) : F(\Theta(\Pi), \Pi) = 0 \text{ for } \Pi \in CI_{1-\alpha}(\Pi)\} \quad (8)$$

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, for any sequence of multinomial distributions satisfying the second condition in the assumption, we have*

$$\lim_{n \rightarrow \infty} Pr\{\Pi \in CI_{1-\alpha}(\Pi) \text{ and } \Theta \in CI_{1-\alpha}(\theta)\} = 1 - \alpha. \quad (9)$$

This theorem can also be shown for any well-behaved functional of θ . For example, if we are interested in $\Delta(\theta) = \theta_1$ (the first element of θ), then in (8) we can use the projections of all Θ to its first component, i.e., $CI_{1-\alpha}(\theta_1) = \{\Delta(\Theta) : \Theta = F(\Pi) \text{ for } \Pi \in CI_{1-\alpha}(\Pi)\}$. This projection will usually be conservative.

Remark 7. *If we assume here that we have a sequence of independent large networks (as opposed to one large network), then there will generally be more scope for applying a limit theorem for independent observations. However, in that setup, one would need to deal with the issue of multiplicity across independent markets. This means that the allocation parameters*

will be network specific unless we assume that the same stable network is being played in every market.

B Details of Simulation Procedures

B.1 Simple Model with One Link (Example 3)

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 function value is zero. We use the simplified QP to confirm results obtained using the more general QP, and to construct identified sets based on different observations (i.e., different vectors of type shares).

The simplified QP is derived from the programming problem as stated in Theorem 3. There are 16 potentially nonzero allocation parameters in the programming problem for this model (see Figure 4), 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 a zero objective function value 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 (e.g., p_{BB}) 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 four 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

Table A1: Equilibrium Type Shares in the Simple Model (Example 3)

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

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 (as a special case of Example 1), 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 5, we fix the structural parameters at the stated values and evaluate a grid of points in the space of admissible vectors of type shares.

B.2 Model with Indirect Connections (Example 2)

B.2.1 Microsimulation Procedure

We generate a number of pairwise stable networks to illustrate the variation in type shares that can arise under this specification. Each simulated network has $n = 500$ individuals, with $n_B = 100$ blacks and $n_W = 400$ whites. To generate a single network, we first draw the vectors of preference shocks for all the individuals in the network. The procedure to find a pairwise stable network then starts with a random initial graph. These initial graphs are generated by independently assigning links with probability $1/(2n)$ and then removing links at random from individuals with more than L links. The success rate of $1/(2n)$ was 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 shocks and move on to draw the next set of preference shocks. We discard any networks that do not converge.

We generated a total of 47 pairwise stable networks in this way (out of 50 total attempts). The degree distribution from all stable 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.

B.2.2 Computational Tools

Our approach requires an enumeration of all possible network types under a given preference structure. Each item in the enumeration is the canonical element from an equivalence class of matrix-vector pairs (A, v) that represent one network type. There can be a large number of distinct (i.e., non-isomorphic) types, so it is useful to have an automated procedure to generate this list. 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 are used to generate preference classes and assess Conditions 1 and 2. All of these objects are constructed prior to the recovery of the identified set, so any computational burden does not directly impact the time it takes to search through the parameter space.

B.2.3 Enumeration of Network Types

To make a list of canonical representations of the network types, we first generate all non-isomorphic adjacency matrices A . This is similar to generating all unlabeled graphs with up to $1 + L \sum_{d=1}^D (L - 1)^{d-1} = 10$ nodes, and various algorithms for graph generation are available.²⁶ For this model, given the limitation on the number of links per node ($L = 3$),

²⁶See, for example, <http://www3.cs.stonybrook.edu/~algorith/files/generating-graphs.shtml> for a list with recommendations.

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 36 non-isomorphic adjacency matrices that are relevant under this preference structure.²⁷

Then to enumerate the different network types we consider all possible combinations of characteristics for 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 description of the types. Also, some types involve symmetries in the local subnetwork, such that it is possible for (A, v) and (A, v') , $v' \neq v$, to represent the same network type. Accordingly we check permutations of the characteristics of the direct alters to ensure that each type in the enumeration is distinct. More general algorithms to test for isomorphisms between graphs with node characteristics (referred to as colors) can also be used.²⁸ In this example there are 356 distinct network types.

B.2.4 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 lists showing 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 (A, v) pair is isomorphic to the result, however, because only the utility of the resulting type is needed. Utility is computed as a function of the characteristics of the direct alters, the number of friends of friends, and the number of mutual friends (see equation (5), where the three rows in the expression correspond to these three factors). 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 lists regarding link deletion record only the characteristics of the direct alters and the numbers of indirect and mutual friends in the resulting types.

²⁷This 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.

²⁸For example, the nauty and Traces programs, available at pallini.di.uniroma1.it.

B.2.5 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. 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 that would result for an individual of type t if a link were added to an individual of type s . For any cases where either t or s already has L links, this entry is of course blank. Otherwise, the resulting type is found by: (1) adding a link to an unoccupied row for a direct alter in the adjacency matrix A_t , (2) inserting the characteristic of the ego from type s into that row of the vector v_t , and (3) adding links to indicate any direct alters from 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 same as the canonical representation for that type, however, so we apply an algorithm to test for graph isomorphisms to find this type within the list of canonical representations. We wrote our own simple algorithm, which considers certain permutations of A and v , but again more general algorithms could be used.²⁹

B.2.6 Distribution of Preference Classes

Given a vector of structural preference parameters θ , the distribution of preference classes in the population is approximated by Monte Carlo integration. We use $S = 10,000$ independent draws of the preference shock vectors. 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 in this simulation approximates its true probability of occurrence.

This procedure yields a set of preference classes $\mathcal{H}(\theta)$ and their probabilities $\{P_{H|x}(\theta)\}$, where θ indicates the dependence on the structural parameter values. The contents of the set $\mathcal{H}(\theta)$ may change with different values of θ (i.e., which preference classes H appear in $\mathcal{H}(\theta)$ may change, not just their probabilities $P_{H|x}(\theta)$). That is because many preference classes have very low probabilities of occurrence, and so may not be realized with even $S = 10,000$ draws, depending on the values of θ .

²⁹We only need to consider a limited number of permutations of A and v because the canonical representation involves an ordering convention that places 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.

To save memory we completely ignore any types that are not either observed in the data or adjacent to an observed type (via addition or deletion of one link). Then, to give a sense of the magnitudes of these sets, at the true parameter values we generate 249 different preference classes (i.e., $|\mathcal{H}(\theta_o)| = 249$). The number of potentially nonzero allocation parameters is equal to the sum of the cardinalities of all these preference classes: $\sum_{H \in \mathcal{H}(\theta_o)} |H| = 5,013$. If we do not remove the unobserved and non-adjacent types, there are 278 preference classes with a total of 12,812 potentially nonzero allocation parameters.

B.2.7 Construction of QP Objective Matrix

Section 5 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). Recall that there is one row in S for each potentially nonzero allocation parameter (i.e., each $\alpha_H(t)$ such that $t \in H$, $H \in \mathcal{H}(\theta)$). The nonzero elements of the row in S for allocation parameter $\alpha_H(t)$ indicate all the allocation parameters (e.g., $\alpha_G(s)$) which correspond to any type s such that an individual of type t with preferences in class H would prefer to add a link to someone of type s (who is at a distance greater than $2D$).

Before running the procedure to populate the rows of S , we first concatenate the contents of all the preference classes in $\mathcal{H}(\theta)$ into a long vector. Hence this vector contains a list of network types, each of which corresponds to an allocation parameter (e.g., the t in $\alpha_H(t)$). The length of this vector is the same as the number of rows and columns in S . We then use this vector, along with the matrix defined in B.2.5, to populate the rows of S as follows.

Consider the row of S for an arbitrary allocation parameter $\alpha_H(t)$. Also recall that each entry $[t, s]$ of the matrix from B.2.5 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 B.2.5 and identify any entry whose value \bar{t} is contained in H . The column positions of these entries correspond to the types s that an individual of type t with preferences in class H would prefer to link to. Thus we have a list of desired alter types s . We then identify the elements of the long vector of types described above that match this list. The positions of these matches correspond to all the allocation parameters that, if nonzero, would yield individuals of some type s with whom someone of type t with preferences in class H would prefer to link to. Therefore, in the row of S for allocation parameter $\alpha_H(t)$, the entries at these positions (i.e., columns) are set to one, while the rest are set to zero.

To save memory, S is stored as a sparse binary matrix. Also, because the membership and order of the preference classes contained in $\mathcal{H}(\theta)$ can change with θ (see section B.2.6), the matrices S and Q are re-constructed for each candidate parameter vector θ . Accordingly, going forward we write $Q(\theta)$. 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 .

B.2.8 Quadratic Programming Problem and Solution

The exact formulation of the QP problem used in this exercise is as follows:

$$\min_{\{\alpha_H(t), t \in H\}, \beta^+, \beta^-} \alpha^\top Q(\theta) \alpha$$

subject to:

$$\sum_{t \in H} \alpha_H(t) = 1, \quad \forall H \tag{10}$$

$$0 \leq \alpha_H(t) \leq 1, \quad \forall H, t \tag{11}$$

$$\sum_H P_{H|x}(\theta) \alpha_H(t) + \beta^+(t) - \beta^-(t) = \pi_x(t), \quad \forall t \tag{12}$$

$$0 \leq \beta^+(t) \leq \delta^+(n), \quad \forall t \tag{13}$$

$$0 \leq \beta^-(t) \leq \delta^-(n), \quad \forall t \tag{14}$$

$$\sum_t \beta^+(t) \leq \Delta^+ \tag{15}$$

$$\sum_t \beta^-(t) \leq \Delta^- \tag{16}$$

To speed the solution we first use a linear programming problem to obtain starting values for the allocation parameters, as those which minimize the sum of absolute deviations between the observed and predicted type shares. This LP is similar to the QP above, except for the objective function and the absence of upper bounds on the slack variables. It is specified as

follows:

$$\begin{aligned}
& \min_{\{\alpha_H(t), t \in H\}, \beta^+, \beta^-} \sum_t (\beta^+(t) + \beta^-(t)) \\
& \text{subject to:} \\
& \sum_{t \in H} \alpha_H(t) = 1, \forall H \\
& 0 \leq \alpha_H(t) \leq 1, \forall H, t \\
& \sum_H P_{H|x}(\theta) \alpha_H(t) + \beta^+(t) - \beta^-(t) = \pi_x(t), \forall t \\
& 0 \leq \beta^+(t), \beta^-(t), \forall t
\end{aligned}$$

The values of the slack variables in the solution to this problem, denoted as $b^+(t)$ and $b^-(t)$, are also used to define the limits Δ^+ and Δ^- in constraints (15) and (16) of the QP. Specifically we set the values of Δ^+ and Δ^- equal to $\max\{\frac{1}{2} \sum_t [b^+(t) + b^-(t)], 6/n\}$.³⁰ This limits the sum of absolute errors in the QP to the (optimal) sum of absolute errors from the LP, 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 (13) and (14) of the QP, 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 amounts and the floor of $6/n$ in the formula for Δ^+ and Δ^- can be thought of as tuning parameters.

To solve the QP, we use an active set algorithm in the program KNITRO. This algorithm 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.

³⁰In the actual formulation of the problem we define Δ^+ to be twice this amount, even though the smaller value for Δ^- binds. (The sums of the observed and predicted type shares are both equal to one, which forces the sums of the positive and negative deviations to be equal.) This seems to help the solver converge more quickly. While searching for a solution it violates the constraints to some extent, and it is apparently beneficial to give more headroom for the positive deviations.

B.2.9 Construction of the Identified Set

In concept, the identified set is a level set in the space of structural parameters, where the optimal function 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 to 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 throughout the parameter space.³¹

Given the fact that error is introduced by having data from a finite network sample, we allow for small positive values of the objective function and of the deviations between observed and predicted type shares. Specifically, 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 . 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.³² For the results plotted in Figures 9 and 10, we generated a total of 7,090 such vectors.

B.2.10 Computational Performance

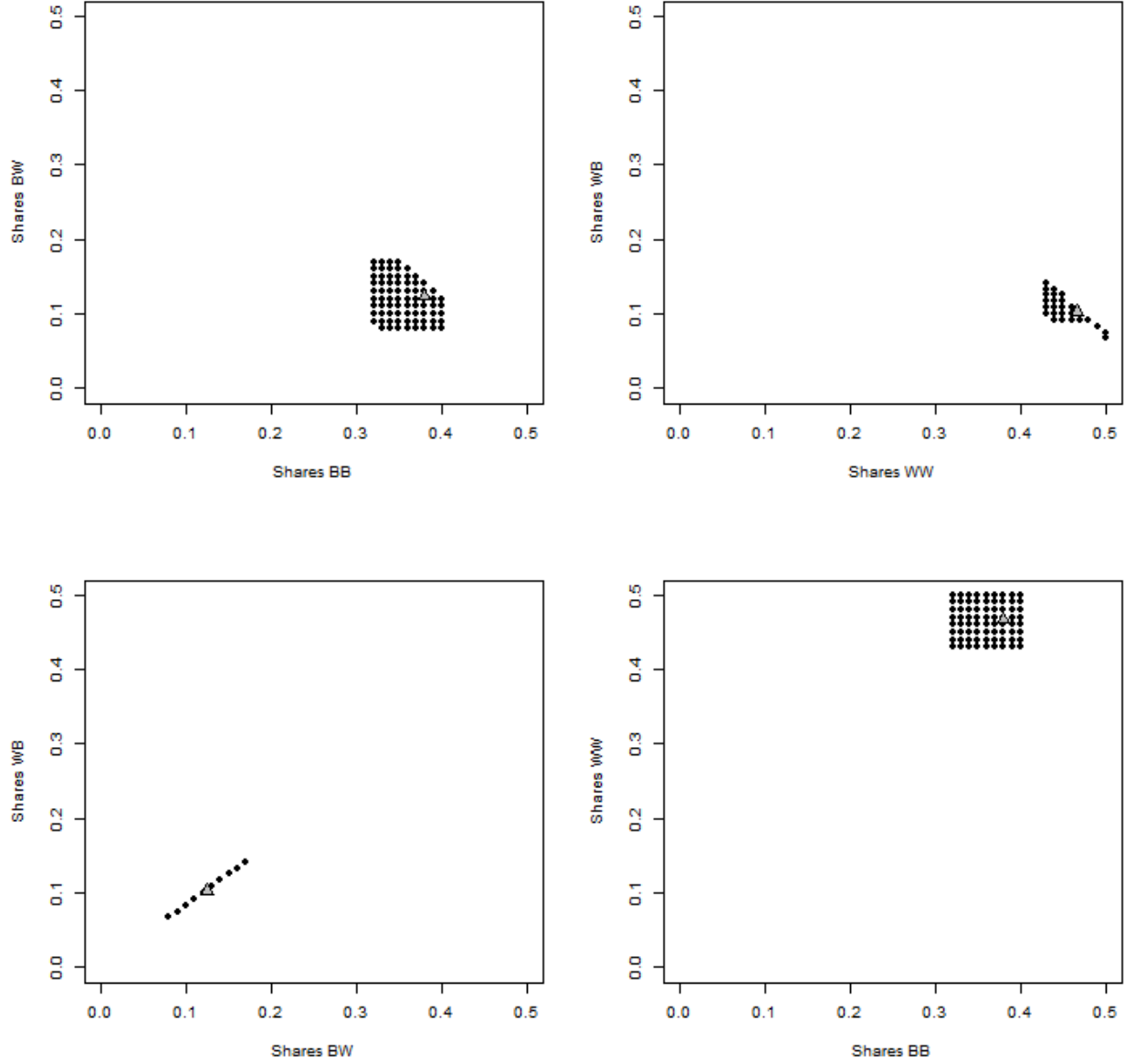
The procedures to generate the identified set were run on machines with Intel® Xeon® 5160 processors (3 GHz base frequency) and 16 GB of physical memory. Computations were not parallelized, except in the “embarrassingly” simple sense that multiple Markov Chains were run on different machines. The time 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 (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.

³¹For the results shown here, we used both the Metropolis-Hastings and slice sampler algorithms in Matlab.

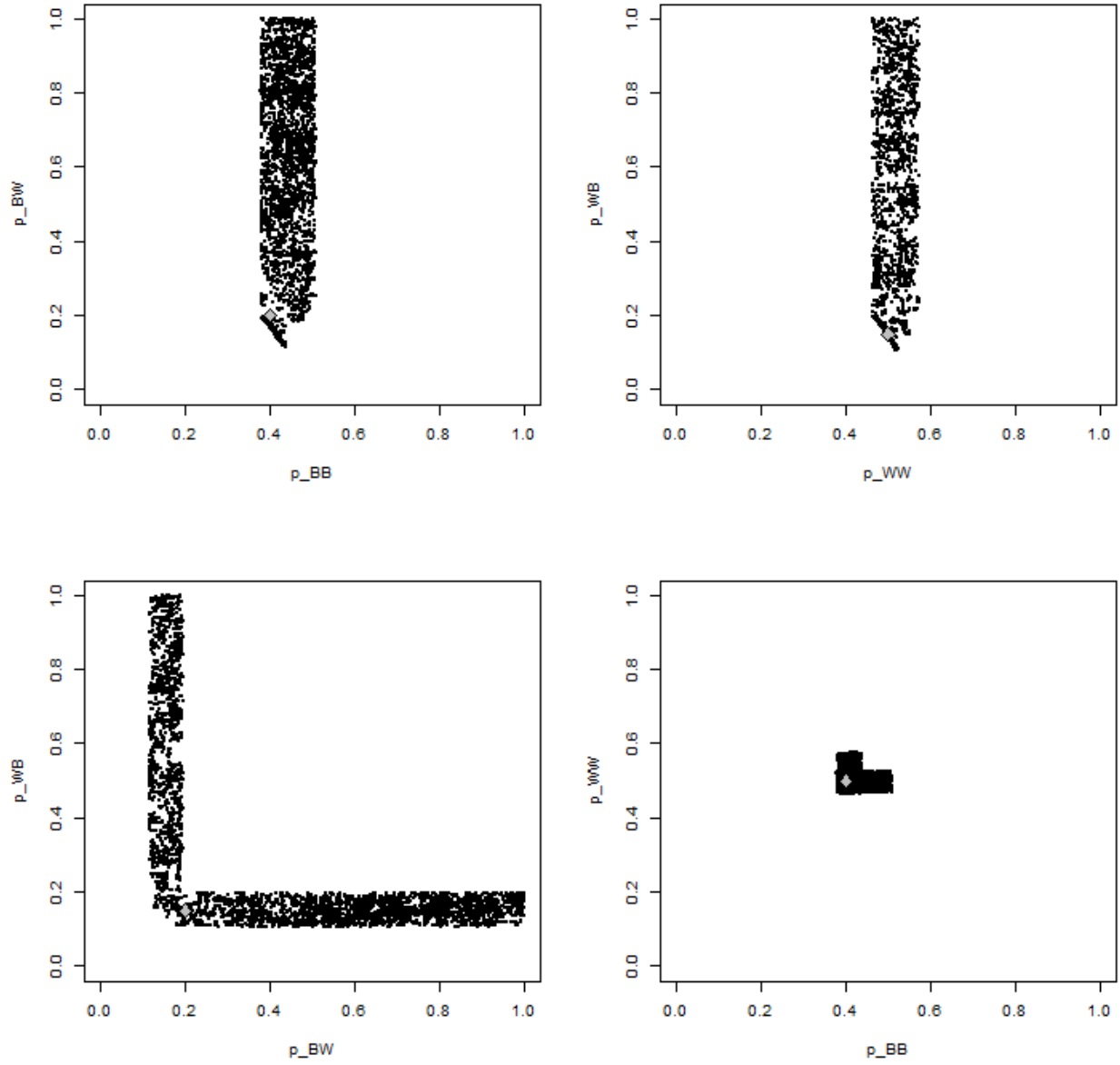
³²The 95% threshold is somewhat arbitrary, but the results are robust to the value that is used. Relatively few parameter vectors produce pseudo-densities that are between 90% and 99% of the maximum value, for example.

Figure 5: Equilibrium Type Shares in Simulation of Example 3



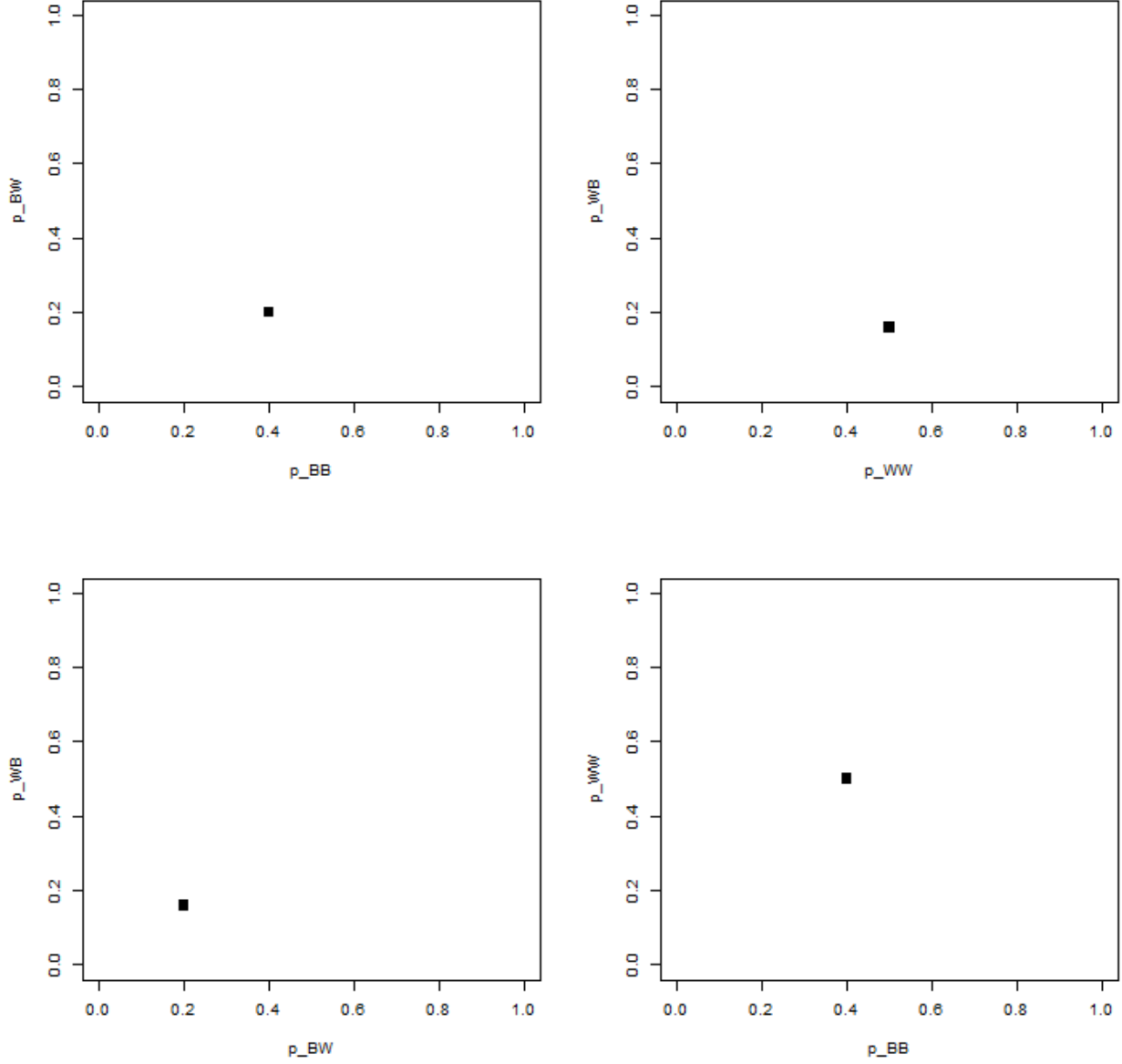
Notes: Points illustrate the full set of equilibrium type shares under parameter values $p_{BB} = 0.40$, $p_{BW} = 0.20$, $p_{WB} = 0.15$, and $p_{WW} = 0.50$. Triangles indicate the type shares from one finite network simulation, with values $\pi_{(B,B)} = 0.38$, $\pi_{(B,W)} = 0.12$, $\pi_{(W,B)} = 0.10$ and $\pi_{(W,W)} = 0.47$ (conditional on race of the ego).

Figure 6: Identified Set from One Finite Network in Example 3



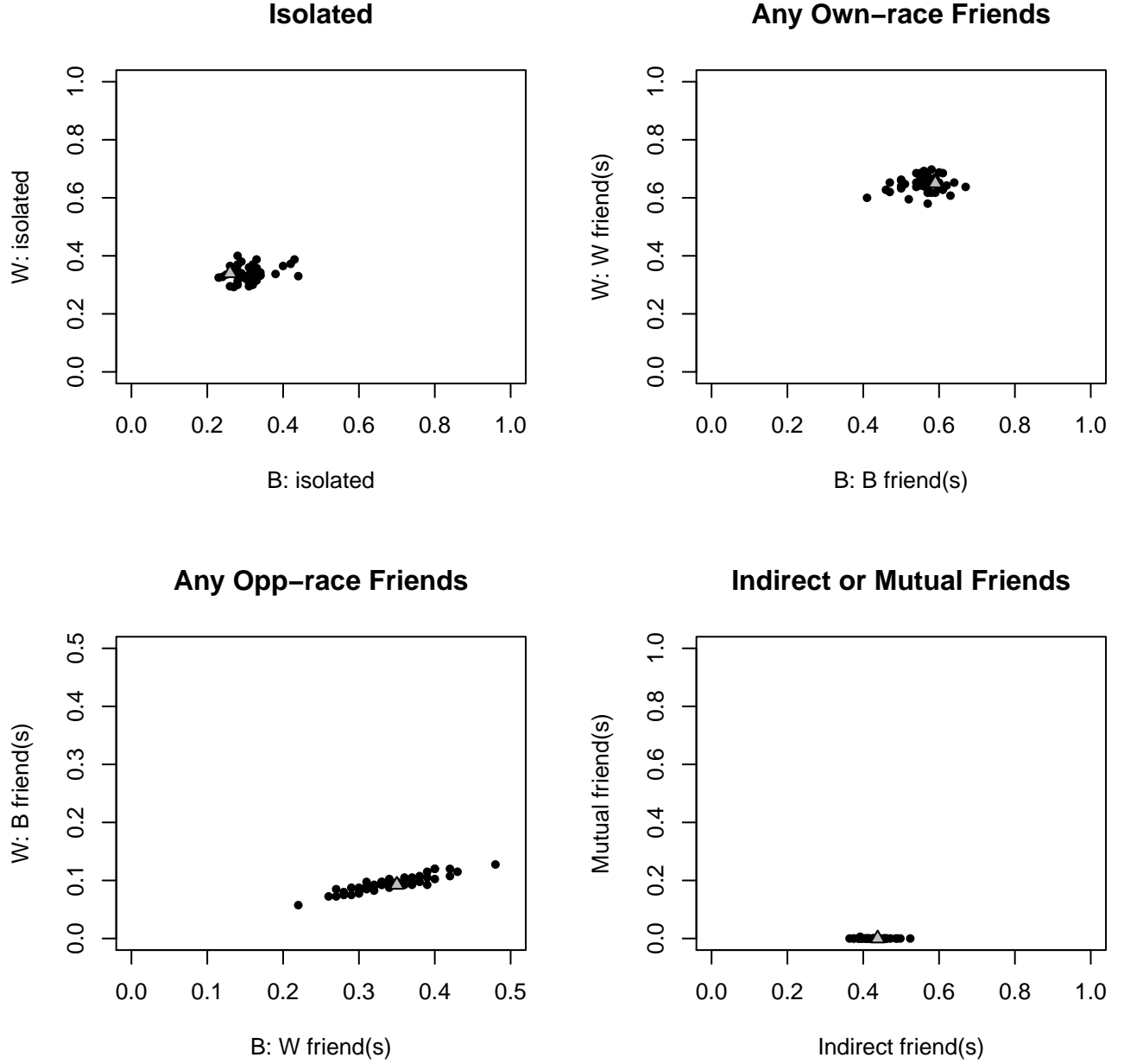
Notes: Points illustrate the identified set obtained using type shares from one finite network (shown with triangles in Figure 5). Diamonds indicate true parameter values: $p_{BB} = 0.40$, $p_{BW} = 0.20$, $p_{WB} = 0.15$, and $p_{WW} = 0.50$.

Figure 7: Identified Set from Four Randomly Selected Networks in Example 3



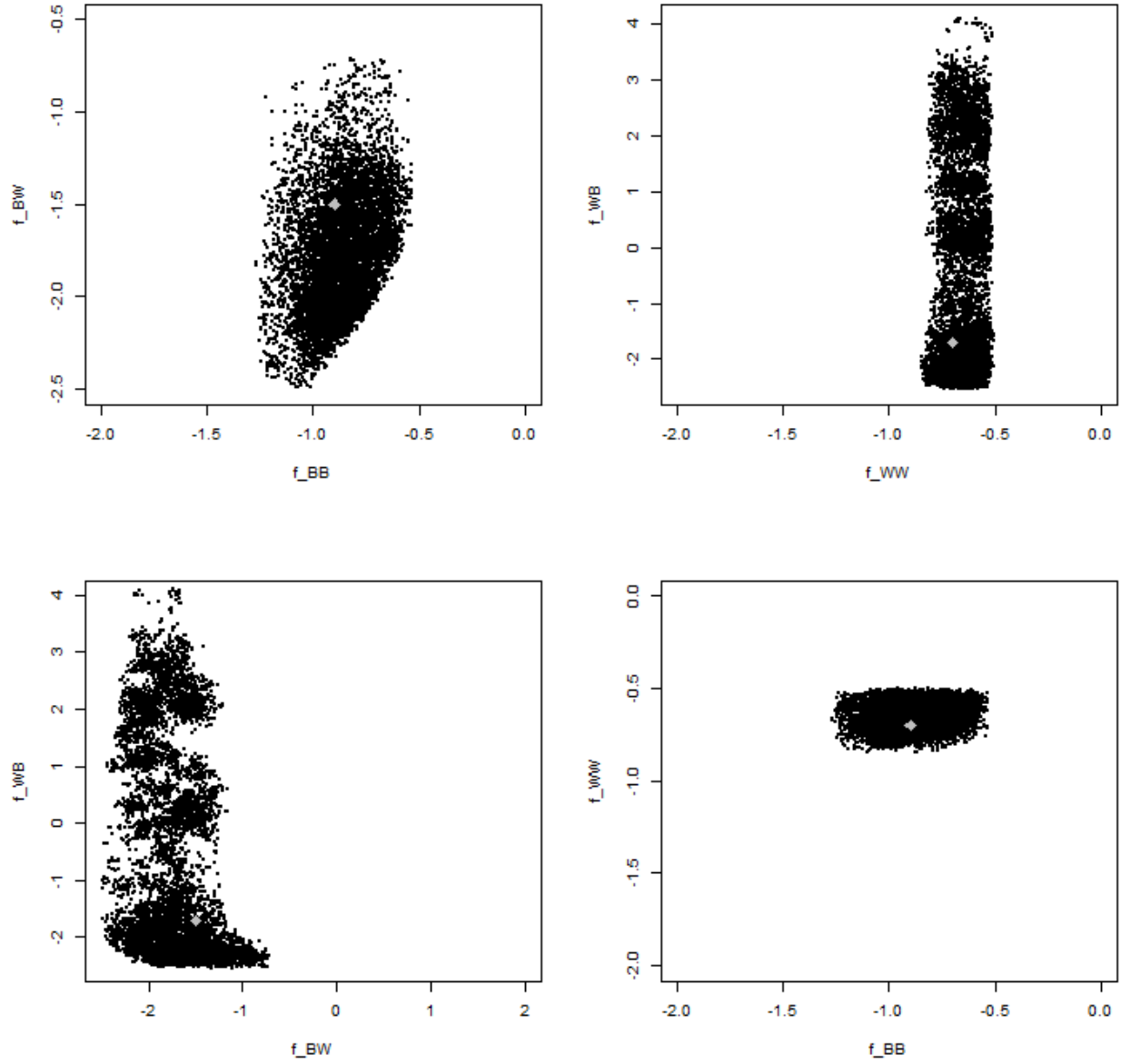
Notes: Blocks illustrate the identified set obtained using four vectors of type shares (shown in Figure A2) that were randomly selected from the full set of equilibrium type shares in Figure 5. A grid of parameter vectors with intervals of size 0.02 in each dimension was evaluated, and the identified set consists of one vector in this grid: $\hat{p}_{BB} = 0.40$, $\hat{p}_{BW} = 0.20$, $\hat{p}_{WB} = 0.16$, and $\hat{p}_{WW} = 0.50$. True parameter values are: $p_{BB} = 0.40$, $p_{BW} = 0.20$, $p_{WB} = 0.15$, and $p_{WW} = 0.50$.

Figure 8: Equilibrium Type Shares in Simulation of Example 2



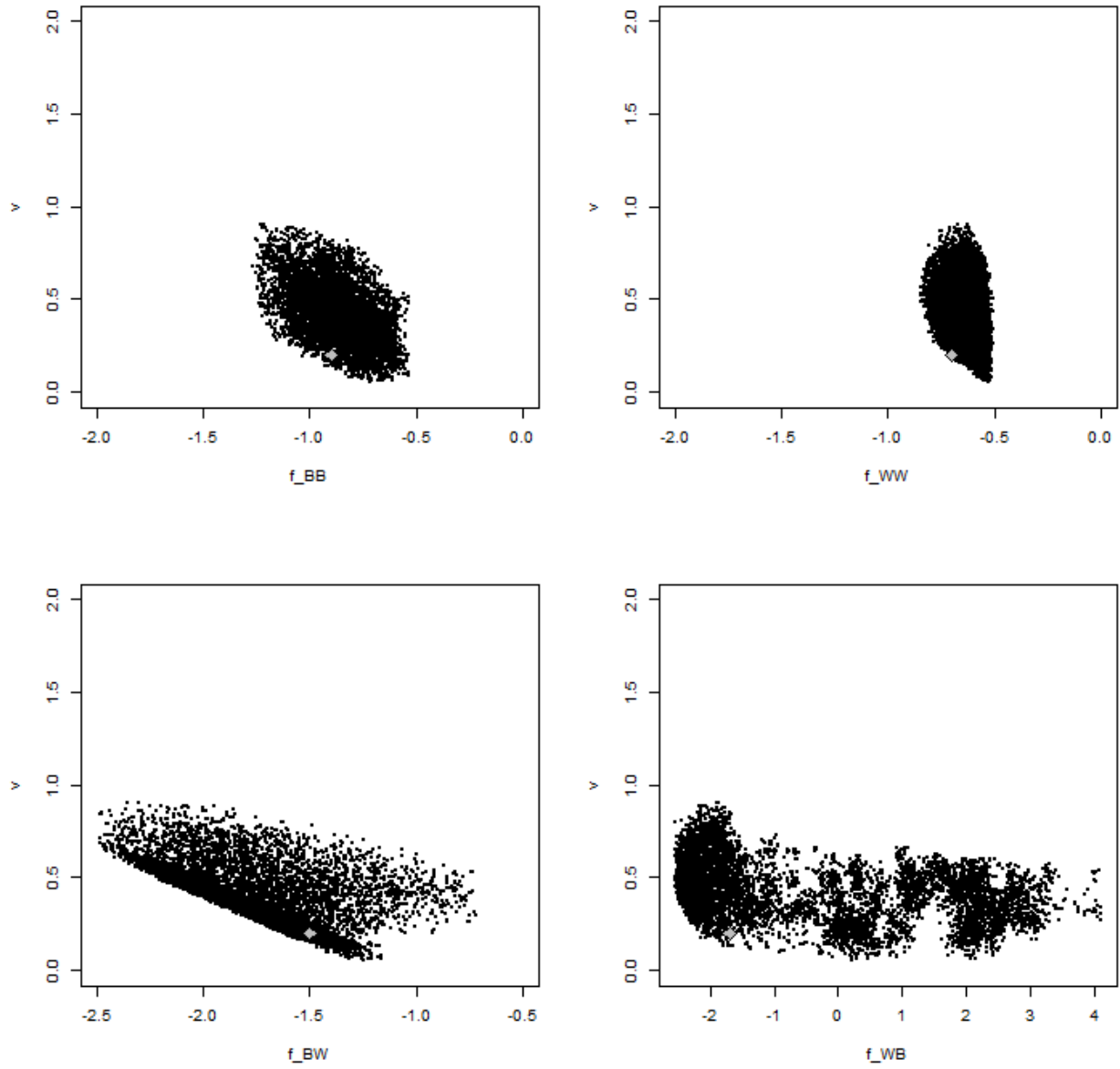
Notes: Figure plots shares of isolated types (x : isolated) and certain combinations of other types: types with any own-race friends (x : x friend(s)), any opposite-race friends (x : y friend(s)), any indirect friends, and any mutual friends. Points represent the shares from different simulated networks,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

Figure 9: Identified Set in Example 2: Values of Direct Connections



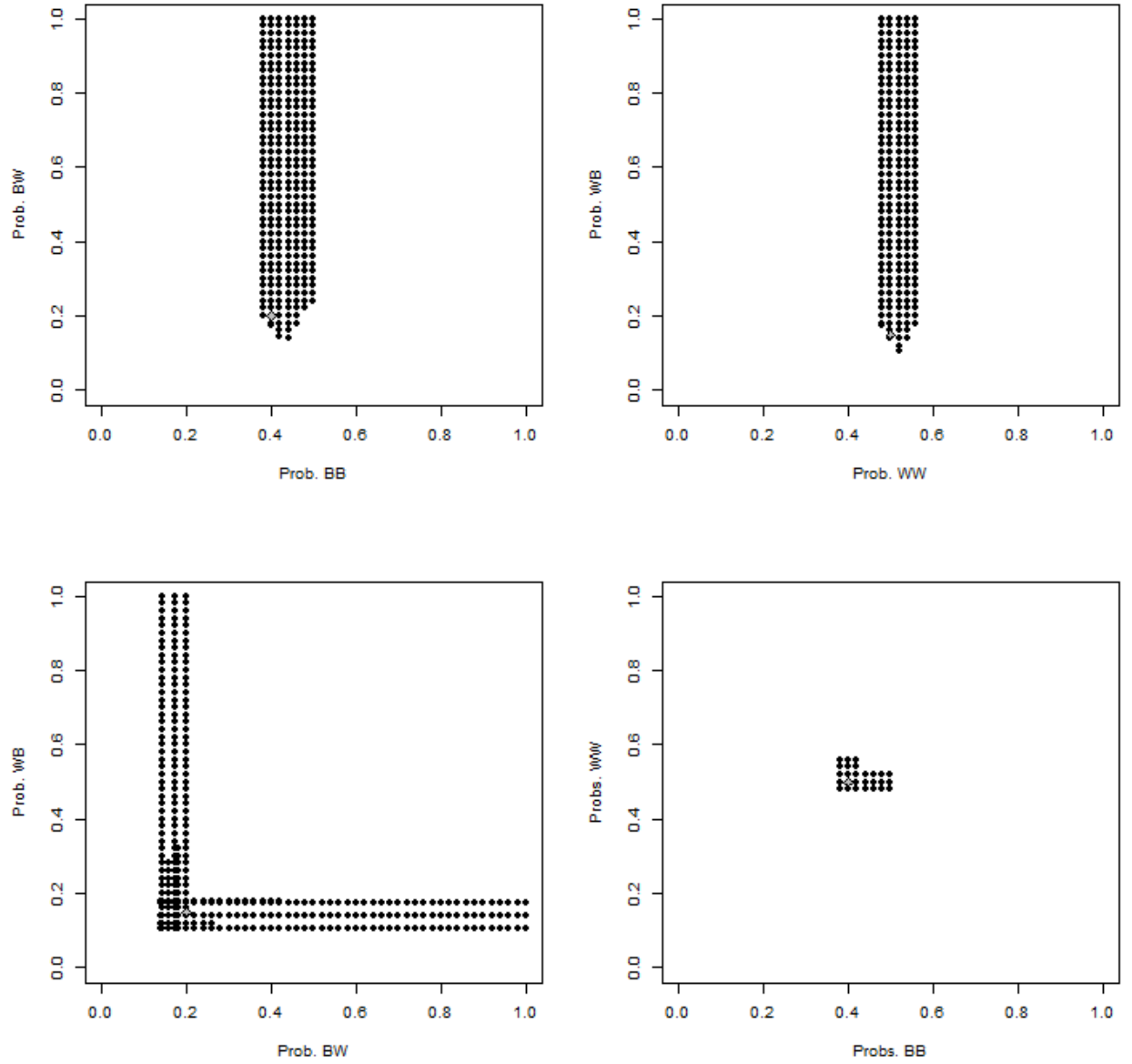
Note: Diamonds indicate true parameter values: $f_{BB} = -0.9$, $f_{BW} = -1.5$, $f_{WB} = -1.7$, $f_{WW} = -0.7$.

Figure 10: Identified Set in Example 2: Values of Indirect vs. Direct Connections



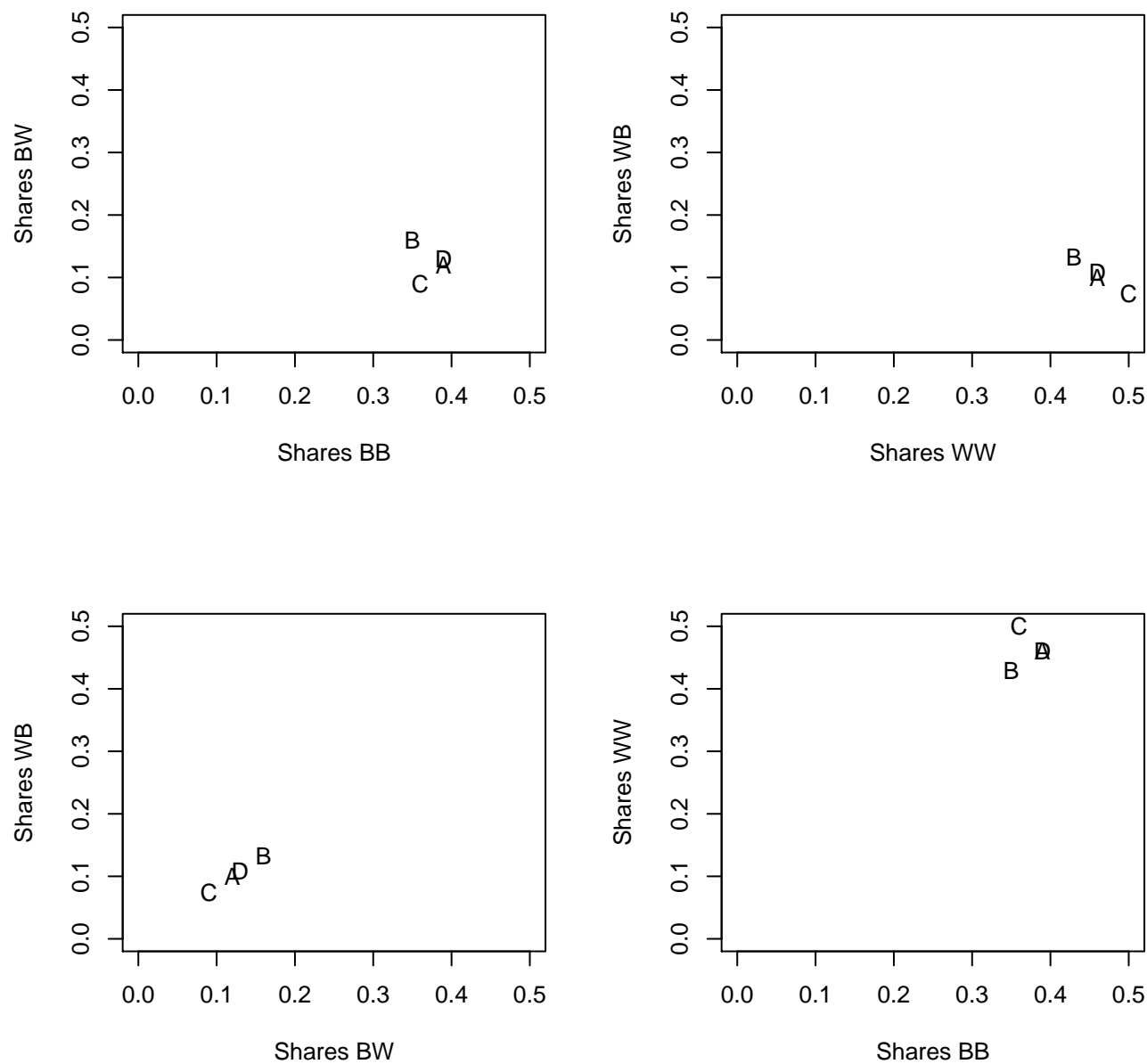
Note: Diamonds indicate true parameter values: $f_{BB} = -0.9$, $f_{BW} = -1.5$, $f_{WB} = -1.7$, $f_{WW} = -0.7$, and $\nu = 0.2$.

Figure A1: Identified Set from One Finite Network in Example 3, Using Simplified QP



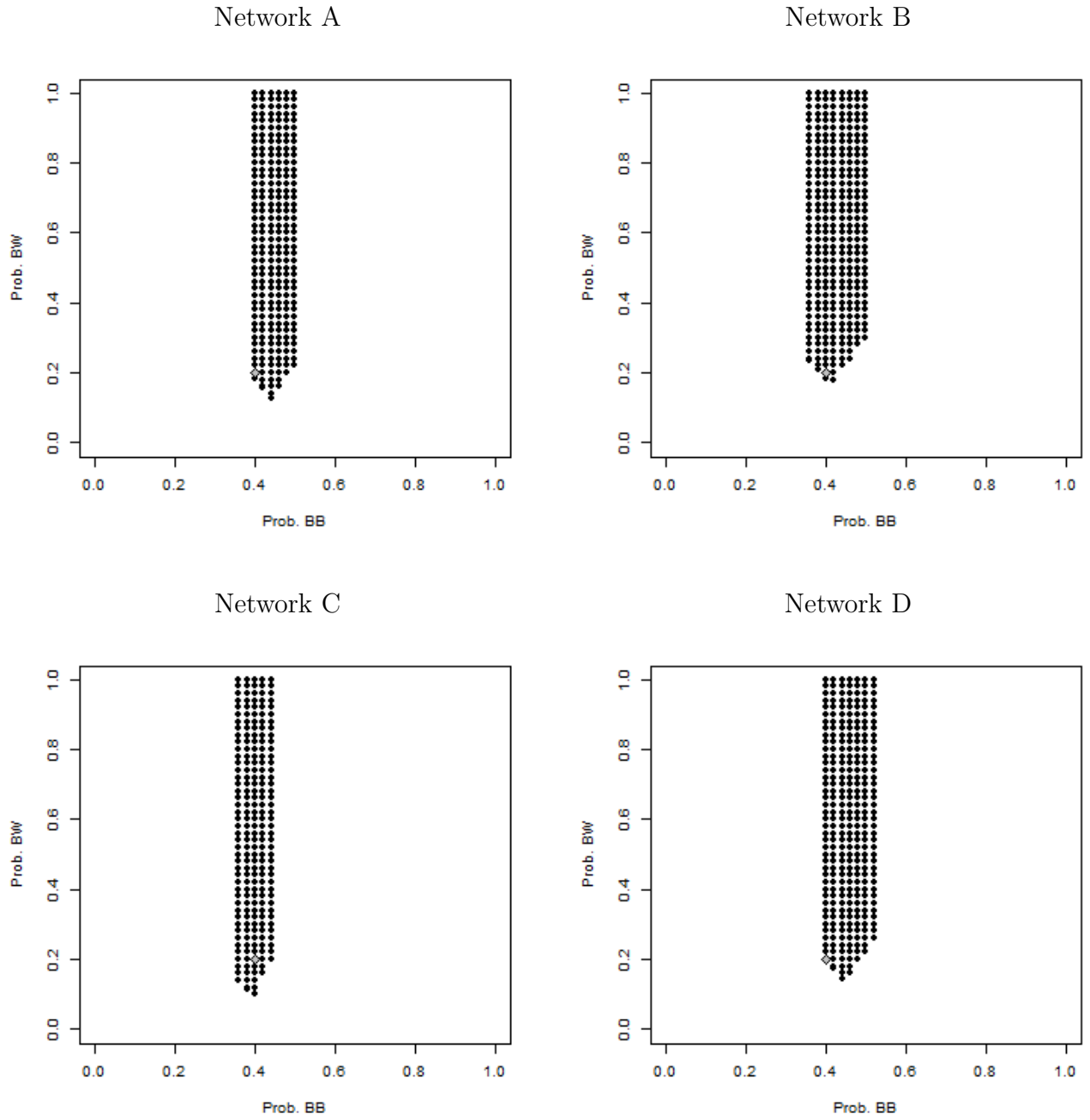
Note: Diamonds indicate true parameter values: $p_{BB} = 0.40$, $p_{BW} = 0.20$, $p_{WB} = 0.15$, and $p_{WW} = 0.50$.

Figure A2: Type Shares in Four Randomly Selected Networks (A, B, C, D) in Example 3



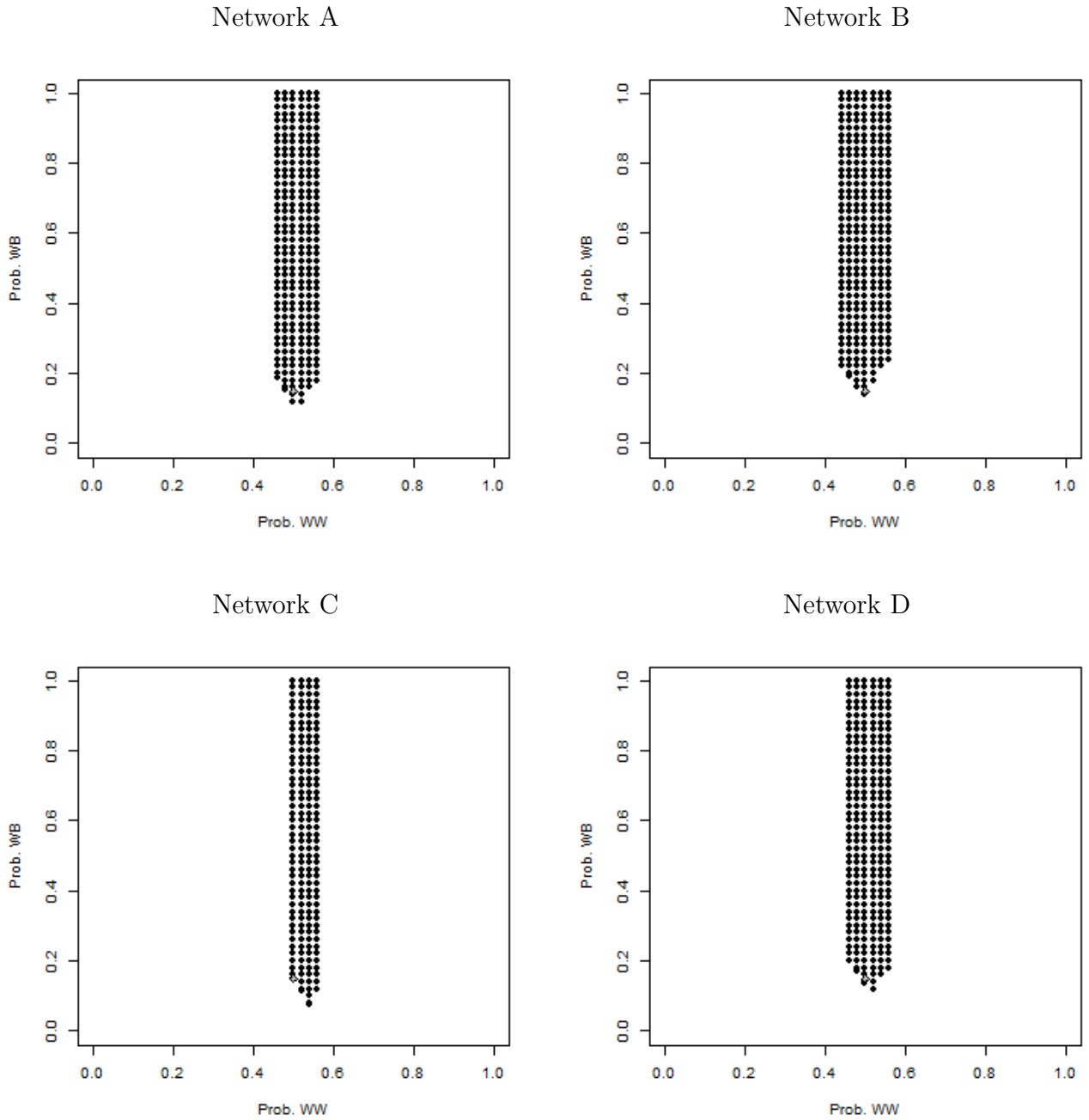
Notes: Letters A, B, C, D correspond to four vectors of type shares that were randomly selected from the full set of equilibrium type shares in Figure 5. Positions of the letters indicate the values of the type shares.

Figure A3: Identified Set from Four Separate Networks in Example 3: Black Preferences



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

Figure A4: Identified Set from Four Separate Networks in Example 3: White Preferences



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

Figure A5: Degree Distribution in Simulation of Example 2 Compared with Add Health Study

