

Identification of Preferences in Network Formation Games ^{*}

Áureo de Paula[†]

UCL, São Paulo School of Economics,
CeMMAP and IFS

Seth Richards-Shubik[‡]

Carnegie Mellon University
and NBER

Elie Tamer[§]

Harvard and Northwestern

This Version: December 2014
Preliminary and Incomplete

Abstract

Given data on a large network, this paper provides a framework for identification of preferences under the assumption of pairwise stability of network links. Network data present difficulties for identification, especially when one allows for links between nodes in a network to be interdependent; e.g., where friends of friends 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 a set of parameters to contain the identified set, under general specifications of preferences. We also provide sufficient conditions—and hence a characterization of the identified set—for two empirically relevant classes of specifications. The paper then provides a quadratic programming algorithm that can be used to construct the identified sets. This algorithm is illustrated in a set of Monte Carlo experiments.

^{*}A version of this paper was presented at the ASSA meetings in Chicago (2011), University of Western Ontario (2013), Stats in Paris: Statistics and Econometrics of Networks (2013), Recent Advances in Set Identification (2013) and the 2nd European Meeting of Network Economics (2014). We thank participants for comments and suggestions and we are grateful to Arun Advani and Cristina Gualdani 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, and IFS. E-mail: apaula@ucl.ac.uk

[‡]Carnegie Mellon University and NBER, E-mail: sethrs@cmu.edu

[§]Harvard and Northwestern, E-mail: ett094@gmail.com

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 a static game of complete information with non-transferable payoffs in which individuals, given a particular utility function, propose links to 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 for this game. 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 dimension of the necessary computations. In fact, a large portion of the computation in our procedure is relatively easily handled through a series of quadratic programming problems.

A network formation model that is founded on a well defined payoff structure for all the players is helpful in determining how outcomes develop given a particular policy or incentive structure. Having an economically founded and estimable model thus becomes a useful ingredient in at least 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. In addition, with an estimable model, we are able to track the effects of various policies or frictions on the kinds and shapes of networks that emerge.

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 model 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 number of mutually exclusive and exhaustive set of local subnetworks that arise in equilibrium. These types are 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* frequency of

types and the model predicted frequencies of types allows us to learn about preferences that are consistent with the observed data. Developing this correspondence in a computationally tractable way represents the main contribution of this paper. This is important because it allows us to construct sets of preferences that are consistent with the data. The main result in the paper provides necessary conditions for a set of parameters to lead to a pairwise stable network with a given predicted distribution of types that matches the data. Also, for certain classes of empirically relevant preference structures, we show that these conditions are sufficient and so are able to characterize the identified set. These results allow us to recover sets of payoff structures that would contain the identified set of preference parameters, and which is the sharp identified set in certain relevant cases.

1.1 Related Literature

In part because of the difficulties indicated previously, the literature on the econometric analysis of network formation models is small, but growing. We focus on a complete information static network formation model without transfers where links are not directed and our solution concept is pairwise stability (see Jackson and Wolinsky (1996)). Other papers differ in one or more ways. Gilleskie and Zhang (2009) and Leung (2014), for example, focus on Nash-type solution concepts and incomplete information. Whereas such models may be of interest in certain contexts, pairwise stability may be more adequate in other circumstances (see, e.g., the discussion in Jackson (2009), p.155). Mele (2010) and Christakis, Fowler, Imbens, and Kalianaraman (2010) rely on a dynamic meeting protocol for the formation of the network, and an extension is found in Badev (2014).¹ Chandrasekhar and Jackson (2014) similarly use a meeting protocol to link certain classes of strategic models to extensions of exponential random graph models. The equilibrium notion involved in such dynamic protocols is different than the static concepts commonly used in the theoretical literature: iid taste shocks are drawn each time individuals meet to contemplate a link, and this chain converges to an ergodic distribution (over networks) that concentrates on equilibrium networks in the absence of the idiosyncratic shocks. Whereas the model estimation targets the invariant distribution for this dynamic protocol in those works, our framework does not rely on a postulated meeting process. This may be desirable especially when dynamic data on

¹A similar model for network determination is also used in Goldsmith-Pinkham and Imbens (2013), which focusses on other outcomes influenced by the network topology.

link formation is not available.²

The payoff structures we analyze are related to those contemplated in Currarini, Jackson, and Pin (2009), Sheng (2014) and Miyauchi (2014). The framework we investigate generalizes the structure in Currarini, Jackson, and Pin (2009), which considers direct connections only. In contrast to Sheng (2014), who suggests to focus on small subsets of players to make inferences on the identified set of parameters, our approach is more scalable since it relies on a large network as its starting point.³ Our proposed method also substantially generalises the models in Sheng (2014) and Miyauchi (2014) as we do not need to rely on the utility restrictions they impose (e.g., no negative externalities) to guarantee equilibrium and facilitate computation. In fact, when an equilibrium does not exist, our prescribed strategy still applies and can produce an empty identified set.

Because matching models essentially aim at characterising 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 (2009), Fox (2010), 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, 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, though not the same as in Jackson and Wolinsky (1996), where only one link at a time is considered.

2 Model Specification and Solution Concept

Our framework considers simultaneous action, complete information games that produce an undirected network. Players announce the set of other players they would like to be connected with, links form if they are mutual, and payoffs are received. Players are $i \in N \equiv [0, \mu]$ where $\mu \in \mathbb{R}^+$ is their total measure. Each 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. Assume that \mathbf{X} and ϵ are independent

²As pointed out in Mele (2010), the meeting protocol acts as an equilibrium selection mechanism, while our approach is agnostic about which equilibrium is observed.

³Miyauchi (2014) considers a model that closely resembles that in Sheng (2014), and uses a common tool in the analysis of super modular games to bound the equilibria and parameters in the model. Miyauchi's approach appears to be feasible up to moderately large networks (e.g., $|N| = 100$).

random vectors.

The network that results from players’ actions is characterized by the adjacency mapping⁴

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

This is a *continuous graph* as there is a continuum of nodes.⁵ Such graphs (particularly a refinement known as *graphons*) are a recent development in mathematics (see Lovasz and Szegedy (2006) and, for a recent review, Lovasz (2012)), and are used as 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 $D = 1$, you do 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

⁴For example, players could make friendship offers ($a_i(j)_{j \in N} \in \{0, 1\}$), and the adjacency mapping would indicate the set of reciprocal positive offers. However these actions do not need to be specified in order to define equilibrium under pairwise stability.

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

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

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 and their characteristics up to any finite distance in the network. 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 at most as many shocks as the potential number of direct connections one can establish and their predetermined characteristics. The number of shocks for each individual then equals $|\mathcal{X}| \times L$, 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 relatively large. Otherwise, if there were i.i.d. preference

number of such links is one, and less than five percent of individuals have more than three links to the same sex (typically considered to be friends rather than romantic partners).

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.

⁸In 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)).

shocks for every potential connection in the network, for example, 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), while in our model, even with large networks, it is always 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.*

This assumption could be relaxed to allow for shocks to each potential payoff-relevant connection, and our main results would still remain. 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 uncorrelation might lead to exceedingly large identified sets.

We showcase next 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). The general specification of the utility function is:*

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

where x is the observable characteristic for individual i , m is the number of direct connections (e.g., friends) with the same 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 set of predetermined characteristics \mathcal{X} is the set of possible individual races. The vector of preference shocks ϵ_i contains one shock for each potential connection to a same-race individual and one 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, the dimension of the preference shocks is reduced further by assuming that the shocks for links with people with any different characteristics are perfectly correlated. We accordingly denote the elements of ϵ_i as $\epsilon_{il}(z)$, where the index 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, so that $\epsilon_i \equiv (\epsilon_{i1}(0) \dots \epsilon_{iL}(0), \epsilon_{i1}(1) \dots \epsilon_{iL}(1))$. 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 you have m same-race friends and n opposite-race friends, you subtract the sums of the first m shocks for same-race alters and the first n shocks for opposite-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). $\sigma \in (0, 1)$ gives diminishing returns. If $\gamma < 1$ there is a preference for same-race friendships, and $\gamma > 1$ 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:*

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

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 there is no double counting of connections here. Some double counting and the positivity of v and w are sometimes used to establish the existence of a pairwise stable network (see, for example, Sheng (2014)). We do not need to impose this latter assumption since, in our case, when the data cannot be reconciled with equilibrium play for a given model, our procedure delivers an empty identified set.

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*

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.

to at least one of the players:

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

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

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 Jackson (2009). As discussed in that reference, an advantage of pairwise stability is that it incorporates in a simple manner the intuition that in a social setting agents are likely to communicate in forming 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.

2.1 Preview of Our Results

In this paper 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, which are formally defined below in Section 3) that would be observed in an equilibrium network. Our main results are to establish that these conditions are necessary for a pairwise stable network to exist with given frequencies of each network type, under any payoff structure that satisfies Assumptions 1 and 2, and are necessary and sufficient for the classes of models in the two examples above. 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 simplified version of Example 1 above. The main role of this example is to illustrate key features of the approach as simply as possible, and we 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). In particular, only direct connections affect utility and individuals have at most one best friend ($D = 1$ and $L = 1$).

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. We observe equilibrium proportions of these types, and the objective is to use these to learn about the $(f_{xy})_{x,y \in \mathcal{X}}$.

To make this operational, 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 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 ϵ 's that determine 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 one of their 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 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 equilibrium 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 be 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 a link to 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. The reason for taking this approach, as we have noted earlier, is the computational burden that would 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, 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) cannot handle the computational problem above.

This paper instead proposes a framework for studying large networks that relies on a particular notion of dimension reduction via anonymity and aggregation. We require, through the specification of preferences, that 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 proportions or measures of these types in the population. The inference question then reduces to collecting all utility parameters that predict proportions of network types that match the proportions estimated in the data.

We now turn to the definition of network types. *Our proposed inference strategy relies on pre-defined “network types” that are descriptive of the local network structure around the individual and derived from the predetermined characteristics X in the population, and whose estimated proportions in the sample are an equilibrium outcome.* The predetermined characteristics are directly observed attributes, such as sex and race, or predetermined behaviors (i.e., 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,” or “an unconnected low income male,” etc., and may include indirect connections as well. Each person in the data belongs to one of these mutually exclusive and exhaustive types or categories, and we assume that the proportion of such network types can be consistently estimated from the observed data.

More precisely, 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

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.

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 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. *Each person belongs to a **Network Type** t . A network type t can be summarized 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 . Given D and L , 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} . Network types are determined in equilibrium and are observed in the data.*

Example 3. *This picks up on the simple example 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 (un-*

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 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 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.

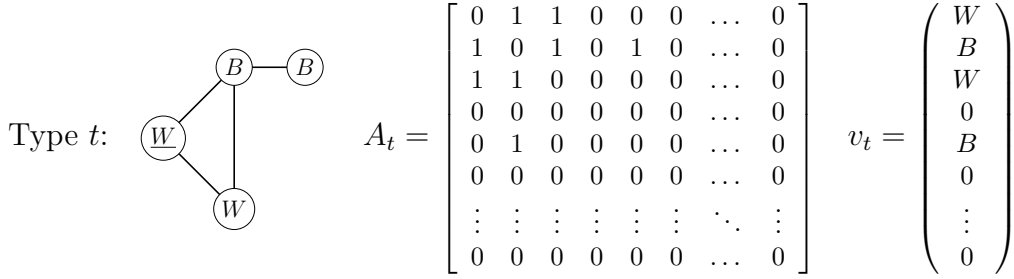


Figure 1: Illustration of a network type, with matrix A and vector v .

linked) 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 so we could just use v in this example.) The data here provide the frequency of the three network types for each race: the proportion of BLACK ALONE, BLACK WITH WHITE FRIEND, BLACK WITH BLACK FRIEND for blacks and similarly for whites. Note here that the types are derived from the specification of preferences and so there is a sense in which types are all that the model “predicts” in terms of observables.

Example 4. Here, we consider a simple case 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 allowed 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 (the characteristic of the ego is underlined) with one black friend and one white friend, who are also friends with each other, and 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 $(f(W, B) + \epsilon_{i1}(B)) + (f(W, W) + \epsilon_{i2}(W)) + \nu + \omega$.

From our first assumption, utility depends only on the local subnetwork that corresponds to an individual’s 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 position of the direct alters within the network type and their predetermined characteristics. The position of an alter corresponds to some row k in the matrix A , and the characteristic of that alter would be v_k . The number of shocks for each individual then equals the number of potential

direct connections multiplied by the cardinality of \mathcal{X} : i.e., $|\mathcal{X}| \times L$. (If shocks to every potential payoff-relevant connection were allowed for, the number of shocks would instead be $|\mathcal{X}| \times L \sum_{d=1}^D (L-1)^{d-1}$ if $L > 1$.) 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

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

Note that this specification 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 1. *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 there is only taste for one friend (and there are no X 's), then there would only be two types, alone and connected. By considering all the possible types that are relevant for a particular class of preferences, there is no loss of information to make inferences about these preferences. Obviously, specifying preferences depends on the empirical question that the model is trying to capture.*

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), preference classes are sets of network types that would satisfy *that individual's* stability requirement for existing links. This corresponds to condition (1) 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

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

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 . However these nodes will have no impact on $u(A_{-l}, v; \epsilon_i)$.

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 $(x, \epsilon) \in \mathcal{X} \times \text{supp}(\epsilon) \mapsto H(x, \epsilon) \subseteq \mathcal{T}$ be the mapping assigning to each pair (x, ϵ) all the network types that satisfy $u(A, v; \epsilon) \geq u(A_{-l}, v; \epsilon)$ where $v_1 = x$ and $l = 1, \dots, L$. A set of types $H \subseteq \mathcal{T}$, or **Preference Class**, is then an element in the range of the mapping $H(\cdot, \cdot)$ as defined above. Preference classes are not observed in the data.

It may be worth noting 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 a connection if assigned to any 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 allocated to network type t_1 or t_2 would not want to drop a link but would do so if allocated to any other network type. This is illustrated in the example below.

Example (3 ctd’). As indicated earlier, network types in this simple network 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 characterize 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 $\epsilon_i(W)$:

- $H_1 = \{(B, 0)\}$: 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, finally, $\{\epsilon \in \text{supp}(\epsilon) : \epsilon_B \geq -f_{BB} \text{ and } \epsilon_W \geq -f_{BW}\}$ for the preference class containing only the isolated network type.

Notice that 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 class 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 $u(A, v; \epsilon) \geq u(A_{-l}, v; \epsilon)$ 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 conditional 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.

Example (3 ctd’). Here, an individual would prefer being best friends with someone of race y over being isolated if and only if $f_{Xiy} + \epsilon_i(y) \geq 0$. 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 $P_{\{(B,0),(B,B)\}|B} = p_{BB}(1 - p_{BW})$ since this is the probability that the person 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 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* that give the proportions of individuals in a particular preference class who 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 ctd'). 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 a white best friend (type (B, W)) would be as follows (multiplied by μ_B):

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

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 a white best friend.

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. 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, in order to be included in the identified set.

4 Identification with Network Types

As noted above, the data allow us to learn the proportions of individuals of each network 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 restrictions are developed using necessary conditions for pairwise stability, which collects a set of (structural) preference parameters that are potentially 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. This result then allows us to learn exactly what sets of preference parameters would be consistent with the equilibrium observed type shares 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 types 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 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,*

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

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. Its purpose 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:

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

where the (local) adjacency matrices $A_{i,-1}$ and $A_{j,-k}$ are obtained by the deletion of the link between i and 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 ctd’). *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 preference parameters to be in the identified set, they must produce preference 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. The condition we use on the allocation parameters 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.¹⁴ Consequently, Conditions (1) and (2) can guarantee pairwise stability in these models and hence the approach would yield the identified 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,

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.

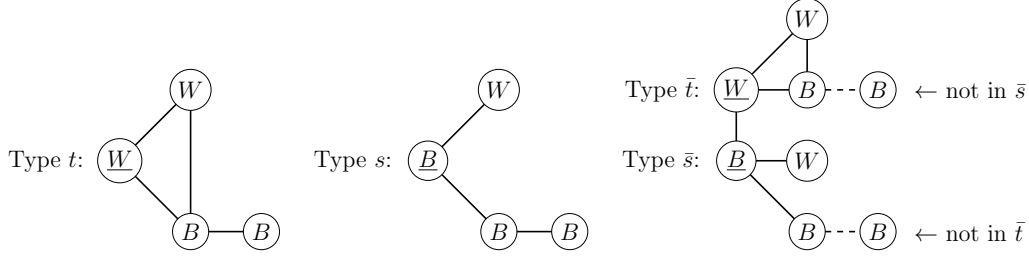


Figure 2: Example of a link added between individuals who are initially distant from each other. *Note:* Dashed lines indicate connections to nodes that appear in only one of the new types, because they are beyond $2D$ from the ego of the other type.

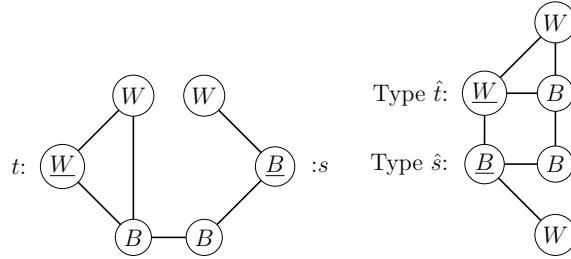


Figure 3: Example of a link added between individuals who are already close in the network

which are based on 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 where the individuals were initially unconnected. These differences would involve loops of length 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, 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 .

Next note that if there is positive measure of type s individuals 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 positive measure of individuals of type s prefer \bar{s} as well, the network is unstable. Accordingly Condition (2) requires that if 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 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 expressed in Condition (2).

Example (3 ctd’). *To apply Condition (2) in our simple model of best-friendships, notice first that only isolated individuals have the capacity to add a link. 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) , 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 a white best friend, or no isolated whites who would prefer to have a black best friend.*

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 , parti-

tion 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). Hence all individuals must be a network type that is within their preference class, 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). 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 there is positive measure of individuals of type t who would prefer to add a link to some individual of type s who is greater than $2D$ away. If so, these individuals of type t would prefer to add a link to any individual of type s who is greater than $2D$ away (because adding a link to any such individual of type s would yield the same marginal payoff). In terms of Condition (2), this corresponds to there being positive measure of individuals of type t who also have \bar{t} in their preference classes ($1_{\bar{t} \in \tilde{H}}$). Because the network is pairwise stable, none of the individuals of type s would prefer to add a link to the individuals of type t who want to link with them. Given that there is positive measure of individuals of type t who want to add such a link, every individual of type s has an individual of type t who wants to link with her (and is greater than $2D$ away). 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 the individual is beyond $2D$), it must be that no individuals of type s would prefer to add a link to any individual of type t who is 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 2. *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 and have finite cardinality of \mathcal{X} , and to have preference shocks pertain to the characteristics of alters rather than their identities. Furthermore one can extend the preference shocks so 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.

In addition, we note that the use of this result for the verification of the identified set can accommodate models where nonexistence is possible. If a particular parameterization cannot generate a pairwise stable network (either with the observed distribution of network types, or with any distribution), then there may be no vector of allocation parameters satisfying Conditions (1) and (2). Hence 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 3. 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. Thus for these empirically relevant classes of models our approach yields the identified set.

To prove sufficiency in these two classes of models we will 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 will 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. We first describe this argument in general, and then provide further requirements and details that are specific to each of the two classes.

To start, fix a parameter vector and, hence, a distribution of preference classes, as well as a distribution of network types. Under these preferences, any network with this distribution of types is unstable. Hence for any such network there must be positive measure of pairs of individuals for whom the presence or absence of a link between them is unstable. To translate this into our conditions, first note that for any network among a set of players N there is a unique vector of allocation parameters that expresses the allocation of the individuals from each preference class to each network type. This is because each individual is associated with one and only one preference class, and one and only one network type.

First we consider existing links ($G(i, j) = 1$). If there is 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 nonexisting links ($G(i, j) = 0$), we first consider individuals who are distant (i.e., $d(i, j; G) > 2D$) from each other in the network. 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 . Therefore the product of the measures given by these expressions is strictly positive, which violates the condition.

Last, we consider nonexisting links between individuals who are $2D$ or less from each other in the network. Here we show that if there is 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 from the above argument. Establishing this involves arguments and requirements that are specific for each example.

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 (i.e., $d(i, j; G) = 2$). So if an individual of type t would prefer to add a link to an (indirect) alter 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 a link to a nearby alter of type s , there is also a positive measure who would prefer to add a link to a distant individual of type s . The same holds for the individuals of type s . If a positive measure of them would prefer to add a link to a nearby alter of type t , there is also a positive measure who would prefer to add a link to a distant individual 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 restriction we can show that having a preference to add a link to a nearby alter ($d \leq 2D$) implies a preference to add a link to someone of the same type at any greater distance. Consider the marginal utility to individual i from adding a link to individual k of some type s . Under the utility specification in Example 2, if k is sufficiently distant in the network, the marginal utility is $f(x_i, x_k) + \epsilon_{il}(x_k) + |N(k)|\nu$, because there is no intersection between i 's existing friends of friends (i.e., $\cup_{l=1}^L, j(l) \neq \emptyset N(j(l))$) and k 's friends ($N(k)$). This holds so long as $d(i, k; G) \geq 4$ because the distance between i 's direct friends and k would be at least 3, so their immediate neighbors ($N(j(l))$ and $N(k)$) could not intersect. Hence the marginal utility of adding a link to any individual of type s is the same so long as that individual is at least 4 steps from i in the existing network G .

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. Compared to adding a link to a more distant individual of type s , the marginal utility would be lower by ν times the number of overlaps between k 's friends and i 's existing friends of friends.

Finally, if $d(i, k; G) = 2$ then the marginal utility of adding a direct link would include the value of mutual friendship (ω), because k would be linked to one or more of i 's existing friends. Specifically the marginal utility would include $\sum_{j(l) \neq \emptyset} G(j(l), k)\omega$ for the new mutual friendship(s). However for each mutual friendship there would also be one fewer friend of friend, compared with the result of adding a link to a distant individual of the same type. (Note that k 's neighbors could be either a mutual friend or a friend of friend to i in the resulting type, but not both.) In addition, k would no longer be a friend of friend. So there would be $\sum_{j(l) \neq \emptyset} G(j(l), k) + 1$ fewer friends of friends in total. Hence the difference in the marginal utility of adding a link to individual k who is at distance 2, compared with adding a link to a distant individual of the same type, would be

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

The maximum number of new mutual friendships ($\sum_{j(l) \neq \emptyset} G(j(l), k)$) 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 individual k who is at distance 2, she would also prefer to add a link to any other individual of the same type at any distance ≥ 4 .

So, as in Example 1, if a positive measure of individuals of some type t would prefer to add a link to a nearby alter of some type s , they would also prefer to add a link to any distant individual 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 a link 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 a link with more distant individuals of the same types. This would violate Condition (2). Combining this with the previous arguments for existing links, and nonexistent 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

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. In particular, the predicted proportions of types must match the observed proportions, and there must be zero allocations from each preference class to any types that are not contained in that class (Condition 1). The latter can be accomplished simply by omitting the allocation parameters that pertain to types not in a given preference class¹⁵.

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 which has one row and column for each allocation parameter (e.g., $\alpha_H(t)$) whose type (t) appears in its 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

¹⁵A similar approach to “solving” the identified set is Honoré and Tamer (2006)’s use of a linear program in a nonlinear panel data model.

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 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 elements of the allocation (i.e., which columns) correspond to any types with whom individuals of type t with preferences in class H would want to add a link (if those alters are greater than $2D$ away). These would be any types s such that a link with an individual of type s (who is greater than $2D$ away) would transform an individual of type t to some type $\bar{t} \in H$. 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 type allocation would prefer to add a link 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 a link to individuals of type t . Thus the Hadamard product of S and S^\top tells us whether an additional link 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 elements of the allocation. (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 ctd'). *To illustrate the objective matrix Q , consider our simplest example with direct links only ($D = 1, L = 1$) and two possible characteristics ($\mathcal{X} = \{B, W\}$). As*

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

Figure 4: Example of Q

noted previously, in this example types can be defined 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 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, one can then establish that:

Theorem 3. *Given a structural parameter vector θ yielding $P(\cdot)$, a network with type shares $\{\pi_x(t)\}$ satisfies Conditions 1 and 2 if and only if*

$$\begin{aligned} & \min_{\{\alpha_H(t): t \in H\}} \alpha^\top Q \alpha \\ & \text{subject to:} \\ & \sum_{t \in H} \alpha_H(t) = 1, \forall H \\ & \alpha_H(t) \geq 0, \forall t, H \\ & \sum_H P_{H|x} \alpha_H(t) = \pi_x(t), \forall t \end{aligned}$$

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 postulate 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 QP 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 4. *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 . However,*

the configuration above 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 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 matrix in Figure 4), which poses a problem for many standard QP solvers. However one can use a more general constrained nonlinear optimization routine. 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 KNITRO.

6 Simulations

We now present simulations based on the example classes of models discussed previously. The main purpose is to illustrate the performance of the approach, in terms of the parameter sets that are recovered and the computational burden that is involved.

In the current draft we present one simulation based on the simple model of Example 3. Simulations based on more elaborate specifications from Examples 1 and 2 will be developed in a subsequent version. The procedures to be described there will aim at providing some guidance on further details of the implementation, such as how to generate the sets of network types and preference classes, and how to construct the matrix Q .

6.1 Simple Example (“Best Friends”)

This is the model introduced in Example 3 (see also subsection 2.1). 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 be used as the primitives in the distribution of preference classes (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, the QP in Theorem 3 can be modified to find *all* equilibrium vectors

of type shares (i.e., π) for a given vector of structural parameters (see Appendix B.1). This is useful to demonstrate the full range of equilibria that are possible. Figure 5 illustrates 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 shows 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 shows the analogous shares for whites ($\pi_{(W,W)}$ and $\pi_{(W,B)}$). This plot indicates further that there cannot be both low shares of whites linked to whites and whites linked to blacks in equilibrium.¹⁶ In the bottom-left plot we see the linear relationship that must hold between the equilibrium shares of blacks linked to whites and whites linked to blacks, because links are one-to-one; i.e., $\mu_B\pi_{(B,W)} = \mu_W\pi_{(W,B)}$.

To find the identified set of preference probabilities (the primitives) from an observed vector of type shares, we conduct a simple grid search over the space of possible preference probabilities, $(p_{BB}, p_{BW}, p_{WB}, p_{WW}) \in [0, 1]^4$. For each vector of preference probabilities in the grid, we check whether the observed type shares are contained in the set of equilibria obtained with the modified QP. We consider four vectors of type shares for this exercise, selected uniformly at random from the set in Figure 5. These vectors are shown in Figure 6. They are labeled as networks A through D, and in each plot the positions of these letters indicate the values of the relevant type shares. For example in the top-right plot, the ‘‘C’’ indicates that the share of whites linked to whites in network C is about 0.42 (on the x-axis) and the share of whites linked to blacks is about 0.10 (on the y-axis).

Figure 7 shows the values of p_{BB} and p_{BW} that are contained in the identified sets recovered from each of these networks, considered separately. Each identified set includes the true probabilities, which are shown in black (the sets include the true values of p_{WB} and p_{WW} as well). The identified sets are reasonably small, except for network B where the identified value for the probability that a black would be willing to link with a white (p_{BW}) is essentially unbounded. This occurs because, under some preference probabilities, the type shares in network B can be generated while allocating no blacks in preference class $\{(B, 0), (B, W)\}$ to be type (B, W) .¹⁷

¹⁶This difference relates to the fact that there are more blacks who would only link with whites ($\mu_B\pi_{BW}(1 - \pi_{BB}) = 0.12$) than whites who would only link with blacks ($\mu_W\pi_{WB}(1 - \pi_{WW}) = 0.09$).

¹⁷In other words, the allocation parameter $\alpha_3(B, W) = 0$ (see the enumeration of preference classes and allocation parameters around Figure 4). If we substitute $\alpha_3(B, W) = 0$ into the expressions for the type

If we were to observe all four networks, the identified set would be the intersection of the identified sets from 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, so we can consistently estimate the type shares in each one.) As Figure 8 shows, the identified set from all four networks together is quite small. In fact, in the grid of points only the probability that a black would be willing to link with a white (p_{BW}) has any identified value other than its true value.¹⁸

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 true parameters need to satisfy under a set of reasonable assumptions on model primitives. This is done by matching the distribution of types observed in the data with the predicted distribution of types from our model under our assumptions. We then show that these same 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. 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 effectively define a set of types. These types are observed in the data and present the crucial link that we explore to learn about preference

shares for blacks in Table A1 and simplify, we have

$$\begin{aligned}\pi_{(B,0)} &= 1 - p_{BB} \\ \pi_{(B,B)} &= p_{BB}(1 - \alpha_4(B, W)p_{BW}) \\ \pi_{(B,W)} &= \alpha_4(B, W)p_{BW}p_{BB}.\end{aligned}$$

Because $\alpha_4(B, W)p_{BW}$ only appear together in this system of equations, we can let p_{BW} take a range of values subject only to $0 \leq \alpha_4(B, W), p_{BW} \leq 1$. For solutions to exist with $\alpha_3(B, W) = 0$ (and hence $\alpha_7(W, B) = 1$), it helps to have $\pi_{(B,B)}$ be relatively small because that allows for larger values of $\alpha_4(B, W)p_{BW}$. This drives down $\alpha_3(B, W)$ for a given share of blacks linked to whites ($\pi_{(B,W)}$). Then $\alpha_3(B, W)$ can be zero if the share of blacks linked to whites is relatively small. It also helps to have $\pi_{(W,W)}$ relatively large because that allows the share of whites linked to blacks ($\pi_{(W,B)}$) to remain relatively small while having $\alpha_7(W, B) = 1$.

¹⁸The grid uses intervals of 0.05 in each dimension, so here the identified set spans less than 0.1 in the values of p_{BB} , p_{WB} , and p_{WW} and less than 0.15 in the value of p_{BW} .

parameters.

In this paper, we focused our work on setting up a model of network formation under pairwise stability and studied how one might study 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. (2014): “Social Networks and the Stepping Stone Effect: The Case of Tobacco and Marijuana,” University of Pennsylvania Working Paper.
- 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.
- ECHENIQUE, F., S. LEE, AND M. SHUM (2010): “Aggregate Matchings,” Working Paper, Caltech.
- FOX, J. (2009): “Estimating Matching Games with Transfers,” Working Paper, University of Chicago.
- FOX, J. T. (2010): “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,” Working Paper, Ecole Polytechnique.
- GILLESKIE, D., AND Y. S. ZHANG (2009): “Friendship Formation and Smoking Initiation Among Teens,” Working Paper, University of North Carolina.
- 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.
- 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.
- LEUNG, M. (2014): “Two-Step Estimation of Network Formation Models with Incomplete Information,” Stanford University Working Paper.
- LOVASZ, L. (2012): *Large Networks and Graph Limits*. American Mathematical Society.
- LOVASZ, L., AND B. SZEGEDY (2006): “Limits of Dense Graph Sequences,” *Journal of Combinatorial Theory, Series B*, 96, 933–957.
- MELE, A. (2010): “A Structural Model of Segregation in Social Networks,” Working Paper, University of Illinois at Urbana-Champaign.
- 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 key feature of our setup, which is allowing for interdependencies in links. But, here, and under some assumptions stated below, we 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 that determines types in the underlying structural model. In addition, pairwise stability, which is the strategic link formation concept in the paper, introduces a set of other parameters that are not of immediate of interest but characterize for example the allocation of each preference case to a given network type. These are related to the particular equilibrium that is played. Hence, the total set of parameters that describe the model is θ . On the other hand, the data are informative only on the measure of types, $\Pi(t)^\top = (\pi(t_1), \dots, \pi(t_{|\mathcal{T}|}))$: since the 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 (the 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 (from the data) 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 (without conditioning on 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 quadratic programming problem, i.e., Θ collects all θ 's where $F(\theta, \Pi) = 0$. Heuristically, to obtain a confidence region on Θ , then, 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 iid 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 iid 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, we must confront the issue of multiple equilibria. Otherwise, since it is possible that in each network there is a *different set* of allocation parameters (or different equilibrium being played) which leads to the number of nuisance parameters (the α 's) growing with the number of networks.

A slightly more realistic data design is one in which we observe one large network¹⁹ and so we can think of this network as a draw from the data generating process. This design is

¹⁹Here, we abstract away also from partial observability of the network which is an important limitation. Typically, given a sampling scheme, a subset of nodes/connections are observed and so if we know the sampling scheme (which is possible), it is possible to extract information about the underlying network and hence the type distribution.

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 type t for 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, mainly,

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

for $k = 1, \dots, T$ 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 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 $N \rightarrow \infty$

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

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

Remark 5. *Note that the fact that we maintain a χ^2 limiting distribution is not essential. If we instead maintain that 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. Also, providing lower level conditions that guarantee that the above condition holds are beyond the central contribution of the paper which we view as building an approach to identification in these models.

Given the above assumption, and to build a (frequentist) confidence region for Θ , one can use a projection type method as follows.²⁰ 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

$$CI_{1-\alpha}(\Pi) = \{\Pi \in S^{|\mathcal{T}|} : G(\hat{\Pi}, \Pi) \leq c_{1-\alpha}(\chi_{|\mathcal{T}|}^2)\} \quad (5)$$

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

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

The above confidence region in (5) is standard and collects the set of 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 out our model in terms of the *set* of θ 's using the quadratic programming function $F(\theta, \Pi) = 0$. The collection of these sets would be a confidence region for the identified set:

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

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 $F(\theta, \Pi) = 0$. So, then we can easily show the following Theorem.

²⁰Alternative 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.

Theorem 4. *Let Assumption 4 above hold. Then, for any sequence of multinomial distributions satisfying the second condition in the Assumption above, we have*

$$\lim_{n \rightarrow \infty} Pr\{\Pi \in CI_{1-\alpha}(\Pi) \text{ and } \Theta \in CI_{1-\alpha}(\theta)\} = 1 - \alpha \quad (7)$$

This theorem can also be shown for any kind functional of θ . For example, if we are interested in $\Delta(\theta) = \theta_1$ (the first element of θ), then in (6) can be 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 6. *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 number of preference 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 (Example 3)

To find the set of all equilibrium type shares, we modify the QP in Theorem 3 by replacing the constraint matching the observed type shares (i.e., $\sum_H P_{H|x} \alpha_H(t) = \pi_x(t)$) with a constraint requiring that the measure of blacks linked to whites must equal the measure of whites linked to blacks:

$$\mu_B \sum_H P_{H|B} \alpha_H(B, W) = \mu_W \sum_H P_{H|W} \alpha_H(W, B).$$

Along with the other two constraints, this ensures the feasibility of the type shares derived from the allocation parameters. Then, because Conditions (1) and (2) are necessary and sufficient for pairwise stability in this model (as a special case of Example 2), we know that all the solutions to the modified QP are indeed equilibria. Any vector of type shares obtained from a solution to the modified QP (minimized at zero) is thus feasible and pairwise stable.

In addition this QP simplifies substantially. Of the 16 allocation parameters that pertain to elements of a preference class (i.e., $\{\alpha_H(t) : t \in H\}$), eight can be fixed or eliminated using

Table A1: Proportions of Types in Equilibrium

| 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) | $\pi_{(W,W)} = \alpha_7(W, B)(1 - p_{WW})p_{WB} + \alpha_8(W, B)p_{WW}p_{WB}$ |

the constraint $\sum_{t \in H} \alpha_H(t) = 1$. Minimization of the objective function requires that four additional allocation parameters are always set to zero: those with a 1 in the diagonal of their corresponding row in the matrix Q (see Figure 4). Positive values of these parameters would assign individuals to be isolated even though they would prefer to be linked to someone in the same preference class (e.g., individuals with preferences $\{(B, 0), (B, B)\}$ assigned to be type $(B, 0)$, who would prefer to be (B, B)). There cannot be a positive measure of such individuals in equilibrium because mutually beneficial links could be added between them. This leaves four allocation parameters: $\alpha_3(B, W)$, $\alpha_4(B, W)$, $\alpha_7(W, B)$, $\alpha_8(W, B)$.

Expressions for the equilibrium type shares as a function of the structural parameters and the remaining four allocation parameters are listed in Table A1. Only two of these allocation parameters are actually free, because either $\alpha_3(B, W)$ or $\alpha_7(W, B)$ must be equal to 1. Otherwise the objective function would not be minimized, as there would be mutually beneficial links to add across race. The other of these two parameters would then be fixed given values of $\alpha_4(B, W)$ and $\alpha_8(W, B)$, using the above constraint for the equality of measures of cross-race linkages.

Figure 5: Equilibrium Type Shares in the Simulation Example

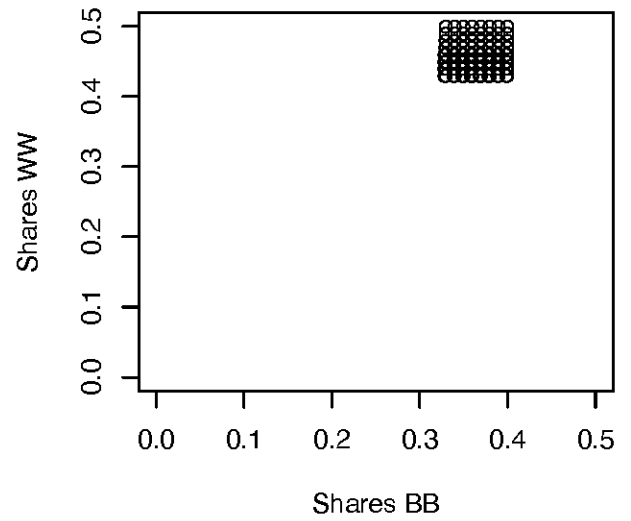
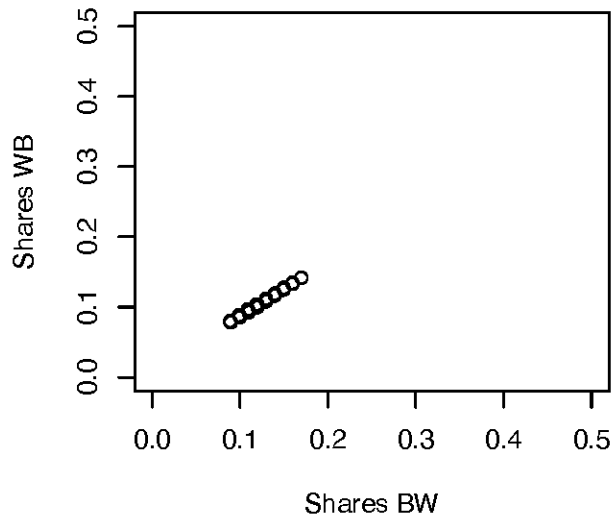
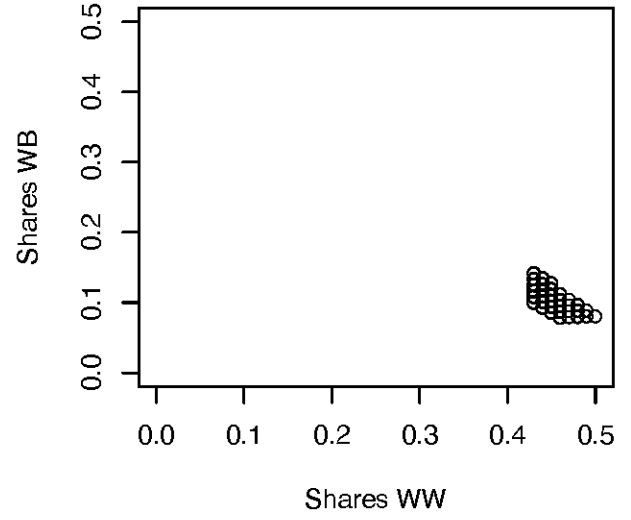
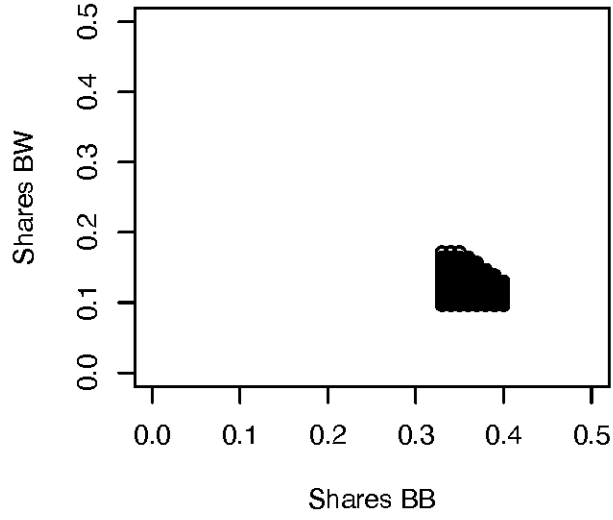


Figure 6: Type Shares in Four Selected Networks (A, B, C, D)

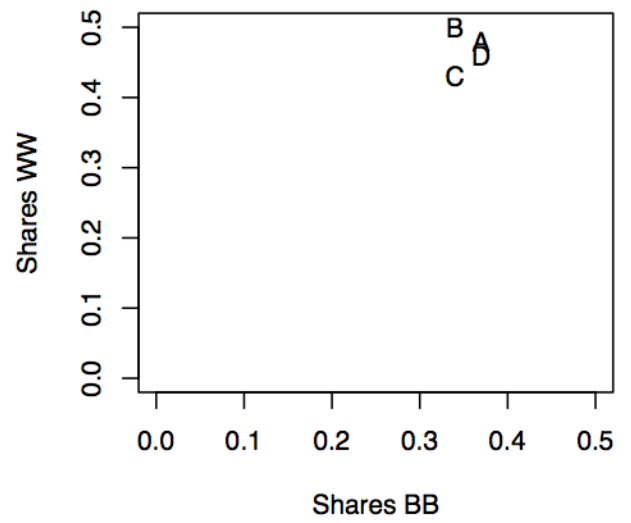
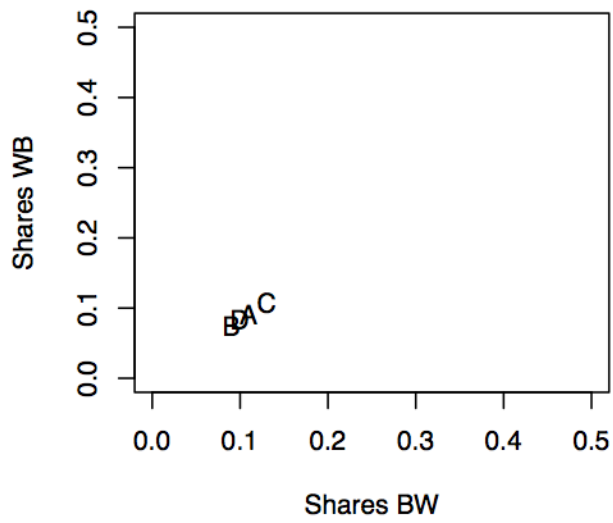
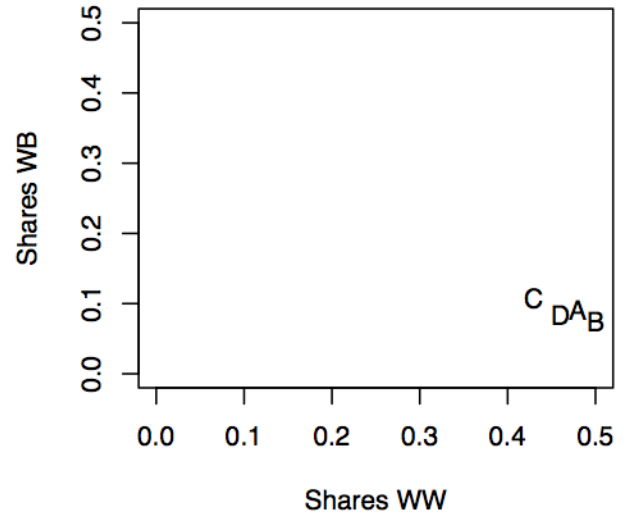
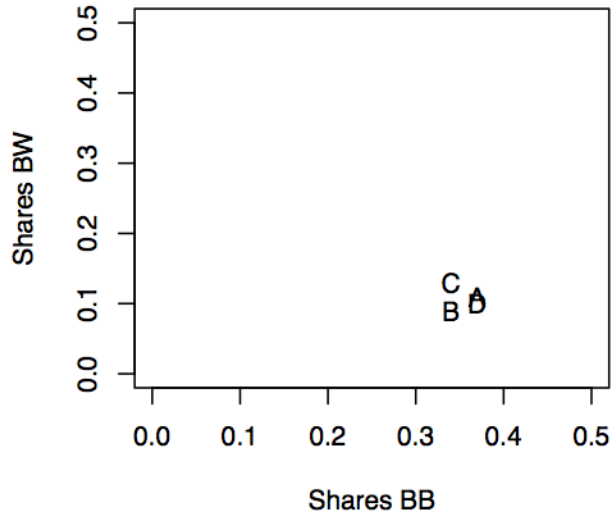


Figure 7: Identified Preferences from Each Network
(*preference probabilities for blacks, true values in black*)

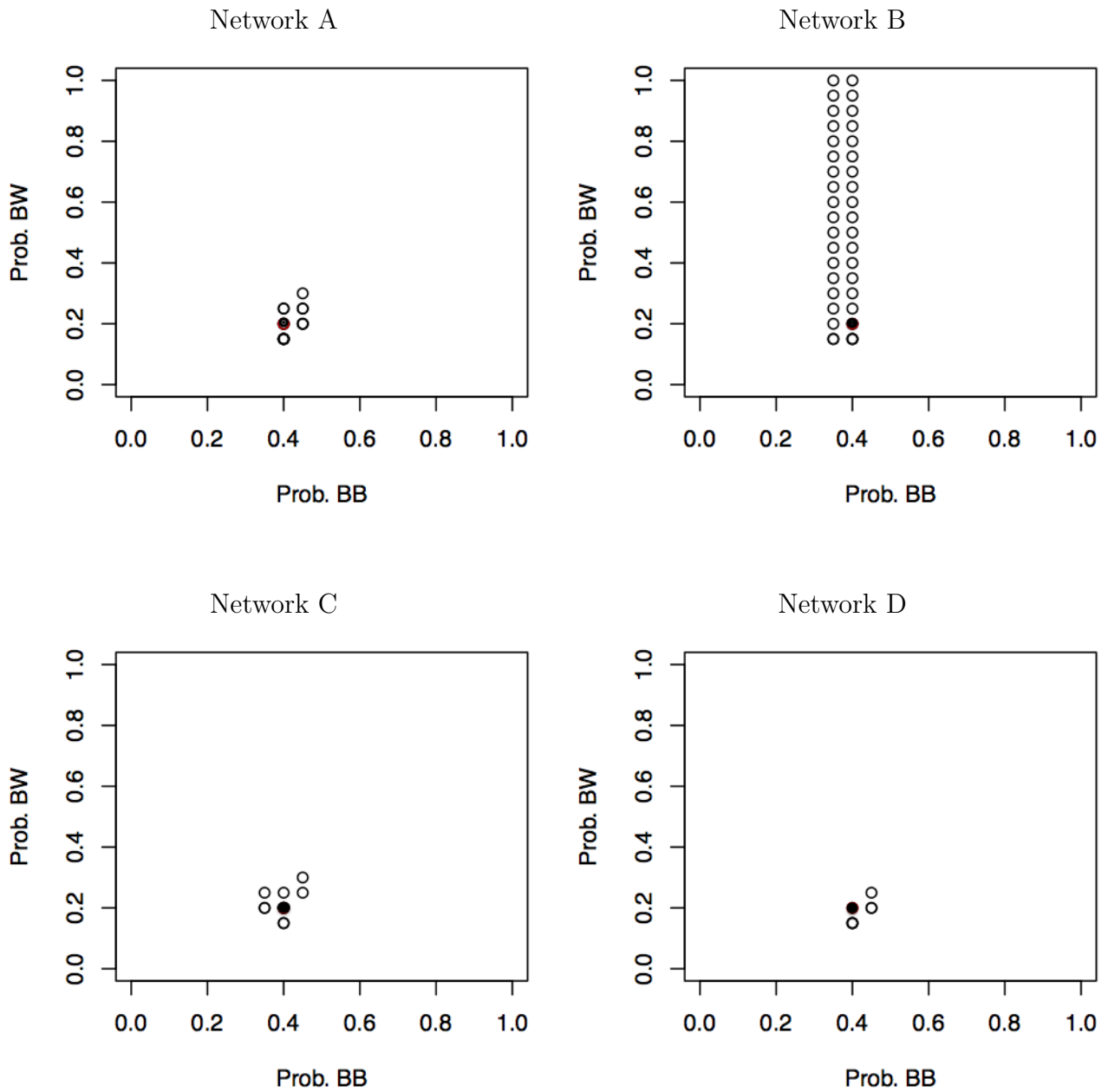


Figure 8: Identified Set from All Four Networks
(*preference probabilities for both races, true values in black*)

