Equilibrium Computation in Discrete Network Games*

Michael P. Leung†

June 28, 2020

Abstract. Counterfactual policy evaluation often requires computation of game-theoretic equilibria. We provide new algorithms for computing pure-strategy Nash equilibria of games on networks with finite action spaces. The algorithms exploit the fact that many agents may be endowed with types such that a particular action is a dominant strategy. These agents can be used to partition the network into smaller subgames whose equilibrium sets may be more feasible to compute. We provide bounds on the complexity of our algorithms for models obeying certain restrictions on the strength of strategic interactions. These restrictions are analogous to the assumption in the widely used linear-in-means model of social interactions that the magnitude of the endogenous peer effect is bounded below one. For these models, our algorithms have complexity $O_p(n^c)$, where the randomness is with respect to the data-generating process, n is the number of agents, and c depends on the strength of strategic interactions. We also provide algorithms for computing pairwise stable and directed Nash stable networks in network formation games.

JEL Codes: C31, C57, C63, C73

Keywords: multiple equilibria, graphical games, network formation, empirical games

*I thank the editor and referees for useful suggestions that improved the paper. This research is supported by NSF grant SES-1755100. It uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

†Department of Economics, University of Southern California. E-mail: leungm@usc.edu.

1 Introduction

Graphical and network formation games have attracted increasing attention in empirical work. Practical use of these models often requires computing the set of equilibria. This is important for evaluating counterfactual policies, for example, assessing the impact of a subsidy on technology adoption in the presence of social interactions (Bhattacharya et al., 2019) or that of busing programs and other reallocation policies (Mele, 2019). Another use is model estimation, which may require computing equilibria in order to evaluate likelihood or moment functions (Bajari et al., 2010a; Miyauchi, 2016; Soetevent and Kooreman, 2007; Xu and Lee, 2015). Also, from a theoretical standpoint, the size of the equilibrium set and its ease of computation are important for assessing the predictive power of a solution concept and its empirical plausibility (Jackson, 2010, Ch. 9.3).

Computing the entire set of equilibria is a difficult problem. A naive brute-force search that checks the equilibrium conditions for every agent is computationally infeasible because the number of action profiles is exponential in the number of agents. Polynomial-time algorithms for computing all equilibria in graphical games are only available for tree networks (Daskalakis and Papadimitriou, 2006; Kearns, 2007). Beyond trees, the "strong conjecture" is that no polynomial-time algorithm exists (Jackson, 2010, Ch. 9.3). For games of strategic complements, the extremal equilibria can be computed in polynomial time, but finding all equilibria between the extremes generally requires exhaustive search (e.g. Jia, 2008).

We provide new algorithms for computing the set of equilibria in graphical and network formation games with finite action spaces. Our algorithms exploit the fact that many agents may have payoff functions such that a particular action is a dominant strategy, that is, always optimal regardless of the actions of other players. For example, in a binary game, an agent endowed with a large enough random-utility shock may find it a dominant strategy to choose action 1. Our algorithm first partitions the network into smaller disjoint subgames on subnetworks whose "boundaries" consist of these agents. The set of equilibria then consists of the "Cartesian product" of equilibrium sets for each of these subgames, which may be feasible to compute using existing methods on account of their smaller sizes.

We provide probabilistic bounds on the complexity of our algorithms for a class of games satisfying a restriction on the strength of strategic interactions. For this

class, we prove that our algorithms terminate in $O_p(n^c)$ evaluations of the payoff function. The value of c is increasing in the strength of strategic interactions, which reveals an interesting trade-off between computability and the economic significance of social interactions.[1] Simulation evidence shows that when strategic interactions are too strong, so that our assumptions are violated, we can expect large, potentially exponential, runtimes.

This class of games with restricted strategic interactions is empirically relevant. In numerical illustrations, we show that our algorithms feasibly compute the set of equilibria of social interactions models estimated by Card and Giuliano (2013) and Xu (2018). Their models obey our restrictions, and the magnitude of the peer effects estimates are economically meaningful. We also note that our restrictions are analogous to the assumption in the widely used linear-in-means model of social interactions that the endogenous peer effect is bounded below one in absolute value (Bramoullé et al., 2009; Calvó-Armengol et al., 2009). Similar restrictions are imposed on autoregressive models in time series and spatial statistics for weak dependence. Furthermore, to our knowledge, the class of models that we study is the only class of static games of complete information for which large-network CLTs are currently available (Leung and Moon, 2019; Leung, 2019). CLTs for large networks are required for inference in the typical setting where the econometrician observes a small set of plausibly independent networks.

The key step for deriving our complexity result is to obtain exponential tail bounds on a certain statistic Δ, which is essentially the size of the largest subgame obtained after partitioning the network according to our algorithm. In fact, Δ corresponds to the size of the largest component of a certain random graph. We approximate this by a branching process for which the desired tail bounds can be obtained more easily, which is a common technique in random graph theory (Bollobás and Riordan, 2008). The argument is also used in Leung and Moon (2019) and Leung (2019) to derive primitive conditions for CLTs in network formation games and graphical games, respectively. We generalize the tail bounds of the former paper to a larger class of sparse graphs. In particular, we do not need to impose the assumption that agents are homophilous.

In economics, homotopy methods have been used to compute game-theoretic equi-

[1] Here "$O_p(\cdot)$" is with respect to the randomness of the data-generating process. Our algorithms are deterministic.

libria for generic games; see Herings and Peeters (2010) for a survey. These are applicable to games with continuous action spaces (Judd et al., 2012) and games of incomplete information (Bajari et al., 2010b). For discrete games, homotopy can be used to compute mixed-strategy equilibria, but the computational cost scales poorly with the size of the game. Our paper is instead concerned with computing pure-strategy equilibria for potentially large, discrete games.

The literature on algorithmic game theory that gives rise to the "strong conjecture" mostly focuses on exact or approximate computation of equilibria in graphical games for general payoff functions, and the typical objective is to obtain worst-case bounds on the algorithmic runtime (Daskalakis et al., 2009). Several papers in this literature explore computability under various restrictions on the network structure. We instead primarily utilize restrictions on the payoff functions, in particular on strategic interactions. We also study stochastic games, where payoffs depend on random and heterogeneous types. This allows us to consider algorithmic complexity from an *ex-ante* perspective with respect to the randomness of the network and types and study the "typical" behavior of our algorithms in large games.

Outline. The next section presents our algorithm and complexity result for graphical games with binary action spaces. We also provide extensions to multinomial and ordered choice. We then present numerical illustrations in §3. In the supplemental appendix, §SA.4 provides analogous results for undirected network formation games under the solution concept of pairwise stability and §SA.5 for directed network formation games.

Notation. We represent a network on n agents as an $n \times n$ adjacency matrix, where the ijth entry A_{ij}, referred to as a *potential link*, is an indicator for whether agent i is connected to j. Following the usual convention, we require that $A_{ii} = 0$ for all agents i, meaning that there are no self links. If $\boldsymbol{A}$ is a symmetric matrix, then it represents an *undirected* network. Otherwise, it is a *directed* network. Consider a directed network $\boldsymbol{A}$. A *directed path* from agent i to j is a sequence of distinct agents starting with i and ending with j such that for each k, k' in this sequence, $A_{kk'} = 1$. The *length* of a directed path is the number of links it involves. A *weakly connected component* is a set of agents such that (1) for each pair of agents i, j in this set, there exists a directed path from either i to j or j to i, and (2) there is no larger set of agents with property (1) containing this set.[2] A *strongly connected*

[2]This definition deviates slightly from standard use in referring to a component as the set of

component is similar, except it requires both a directed path from i to j and j to i. For an undirected network, we refer to a strongly connected component simply as a *component*.

2 Graphical Games

Let $\mathcal{N}_n = \{1, \ldots, n\}$ be the set of agents, which are connected through an undirected network $\boldsymbol{A}$. Each agent $i \in \mathcal{N}_n$ is endowed with a type $T_i \in \mathbb{R}^{d_t}$, which is distributed i.i.d. across agents. Let $\boldsymbol{T} = (T_i)_{i=1}^n$ be the type profile. For now, we assume each agent i takes a binary action $Y_i \in \{0, 1\}$. We later extend the results to more than two actions. Let $\boldsymbol{Y} = (Y_i)_{i=1}^n$ be an action profile. For any i, we may partition $\boldsymbol{Y} = (Y_i, Y_{-i})$, $\boldsymbol{T} = (T_i, T_{-i})$, and $\boldsymbol{A} = (A_i, A_{-i})$, where A_i is the ith row of $\boldsymbol{A}$ and A_{-i} the remaining submatrix, and Y_{-i}, T_{-i} are similarly defined. For a given action profile $\boldsymbol{Y}$, agent i's net payoff from choosing action 1 over 0 is

$$U(S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i), \quad \text{where} \quad S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}) \equiv S(Y_{-i}, T_i, T_{-i}, A_i, A_{-i}),$$

for some function $S(\cdot)$ with range $\mathbb{R}^{d_s}$. Strategic interactions enter payoffs through the vector of statistics $S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A})$ due to its dependence on Y_{-i}.

An action profile $\boldsymbol{Y}$ constitutes a *pure-strategy Nash equilibrium* if for every $i \in \mathcal{N}_n$,

$$Y_i = \mathbf{1}\left\{U(S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i) > 0\right\}.^3 \tag{2.1}$$

Let $\mathcal{E}_{NE}(\boldsymbol{T}, \boldsymbol{A}) \subseteq \{0, 1\}^n$ be the set of Nash equilibria, that is, the set of action profiles such that for each $\boldsymbol{Y} \in \mathcal{E}_{NE}(\boldsymbol{T}, \boldsymbol{A})$, the ith component Y_i satisfies (2.1) for all i.

Example 1. A large literature dating back to Granovetter (1978) studies threshold models of behavior, where agent i chooses action 1 if and only if the number or share of neighbors choosing that action exceeds a threshold (Jackson, 2010; Schelling, 1978). This model has been used to study, for example, product adoption and protests

agents, rather than the subnetwork on this set of agents.

[3]The choice of tiebreaking rule here for indifference is not important for the results. It has no material import in typical econometric applications, where the payoff function is additively separable in a continuously distributed stochastic error.

(Godinho de Matos et al., 2014; González, 2017). It corresponds to the payoff function

$$U(S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i) = \beta \frac{\sum_{j \neq i} A_{ij} Y_j}{\sum_{j \neq i} A_{ij}} - \xi(T_i)$$

for the case in which the fraction of adopting neighbors influences own adoption. Here $S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}) = \sum_j A_{ij} Y_j / \sum_j A_{ij}$ is the fraction of neighbors choosing action 1, and $\xi(T_i)/\beta$ is agent i's threshold. This specification is similar to the well-known model of social interactions studied by Brock and Durlauf (2001), which is the discrete choice analog of the Manski (1993) linear-in-means model that is widely used in applied economics to study social interactions. In place of the term multiplying β, we can also consider type-weighted versions of the average action, for example $S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}) = \sum_j A_{ij} \mathbf{1}\{T_j = t\} Y_j / \sum_j A_{ij} \mathbf{1}\{T_j = t\}$, or nonlinear functions of Y_{-i} and A_i such as the minimum or maximum action. Hoxby and Weingarth (2005) motivate the use of these alternative specifications.

In these examples, strategic interactions only operate through network neighbors of the ego i. We next impose this restriction more generally. Let $\mathcal{N}(i) = \{j \in \mathcal{N}_n \colon A_{ij} = 1\}$ the set of agents connected to i. For any set of agents $G \subseteq \mathcal{N}_n$, let $T_G = (T_i)_{i \in G}$, and likewise define Y_G.

Assumption 1 (Local Interactions)**.** *There exists a function $\tilde{S}(\cdot)$ such that for all $n \in \mathbb{N}$ and $i \in \mathcal{N}_n$,*

$$S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}) = \tilde{S}(Y_{\mathcal{N}(i)}, T_i, T_{\mathcal{N}(i)}, A_i).$$

This says that i's payoffs only depend on the outcomes, types, and potential links of agents connected to i. It is distinguished from, for example, aggregate games in which payoffs depend on some aggregate statistic involving the actions of all agents (e.g. Menzel, 2016), to which our results do not apply.

Econometrician's Information. We assume that the econometrician observes the network $\boldsymbol{A}$ and type profile $\boldsymbol{T}$, and given a known payoff function $U(\cdot)$, her objective is to compute $\mathcal{E}_{NE}(\boldsymbol{T}, \boldsymbol{A})$. Now, typically in practice, $\boldsymbol{A}$ is observed in the data, but types are instead partitioned into an observed and unobserved component $T_i = (X_i, \varepsilon_i)$. Also, both $U(\cdot)$ and the distribution of ε_i given X_i are specified up to some unknown parameter θ, in which case $\boldsymbol{T}$ and $U(\cdot)$ are not entirely known. However,

for counterfactual exercises, a candidate value of θ is typically selected (for example it may be estimated or chosen from an identified set), resulting in a known $U(\cdot)$. Then, for every agent i, the unobserved component ε_i is drawn from the specified conditional distribution under the candidate θ, resulting in an observed type profile $\boldsymbol{T}$.

Network Formation Model. We consider a nonparametric, stochastic model of network formation for $\boldsymbol{A}$. The implementation of the algorithm will not depend on this model. Instead, its purpose is to derive bounds on the complexity of our algorithm. To simplify the exposition, we initially consider a model with no strategic interactions, but we later generalize the main result to the larger class of models introduced in §SA.4, which do allow for strategic interactions.

Endow each pair of agents $\{i, j\} \subseteq \mathcal{N}_n$ with a random utility shock ζ_{ij}, which is i.i.d. across pairs. For all $i, j \in \mathcal{N}_n$ with $i \neq j$, potential links in $\boldsymbol{A}$ satisfy

$$A_{ij} = g_n(\tau_i, \tau_j, \zeta_{ij}), \tag{2.2}$$

where $g_n(\cdot)$ is a $\{0, 1\}$-valued function that varies with the network size n, and $\tau_i \in \mathbb{R}^d$ is a subvector of T_i that influences link formation (and possibly also outcomes). For example, in the context of friendship formation, τ_i may include the race and gender of individual i. Since A_{ij} is an undirected network, we assume $g_n(\tau_i, \tau_j, \zeta_{ij}) = g_n(\tau_j, \tau_i, \zeta_{ji})$.

Semiparametric analogs of (2.2) are commonly used to study link formation (e.g. Fafchamps and Gubert, 2007). Graham (2017) studies estimation of the model

$$g_n(\tau_i, \tau_j, \zeta_{ij}) = \mathbf{1}\{h(X_i, X_j)'\beta + \alpha_i + \alpha_j + \zeta_{ij} > 0\},$$

where $\tau_i = (X_i, \alpha_i)$, and only the subvector X_i is observed. The function $h(\cdot)$ allows for homophily in the X_i's. Our model also allows for homophily in the α_i's. In the case where α_i is correlated with some other unobserved component of T_i that enters payoffs $U(\cdot)$, this generates *unobserved homophily*, which induces a network that is endogenous with respect to the unobserved determinants of outcomes Y_i. Model (2.2) also nests stochastic block models and latent space models, which are the subject of a large literature in statistics (e.g. Bickel and Chen, 2009).

2.1 Strategic Neighborhoods

We next state and motivate a key concept used in our algorithm, which is the notion of a strategic neighborhood. We first need several definitions. For any $G \subseteq \mathcal{N}_n$, recall that $T_G = (T_k)_{k \in G}$ and A_G is the submatrix of $\boldsymbol{A}$ containing only the rows and columns of $\boldsymbol{A}$ in G. Then $\mathcal{E}_{NE}(T_G, A_G)$ is the set of Nash equilibria in the game where the set of players is G rather than $\mathcal{N}_n$.

Define the *non-robustness indicator*

$$\mathcal{R}_i^c = \mathbf{1}\big\{ \inf_s U(s, T_i) \leqslant 0 \cap \sup_s U(s, T_i) > 0 \big\}. \tag{2.3}$$

We say that the equilibrium action of agent i is *robust* if $\mathcal{R}_i^c = 0$ and otherwise that it is *non-robust*. When i's action is robust, either $\inf_s U(s, T_k) > 0$ or $\sup_s U(s, T_k) \leqslant 0$. In the former (latter) case, agent k chooses action 1 (0) regardless of her neighbors' outcomes, which only enter k's payoffs through the first argument of $U(\cdot)$. Hence, $\mathcal{R}_i^c = 0$ implies that i's equilibrium action is a dominant strategy for the given realization of her type.

Define a directed network $\boldsymbol{D}$ on $\mathcal{N}_n$ with ijth entry

$$D_{ij} = A_{ij} \mathcal{R}_j^c.$$

This connects an agent i to a neighbor (with respect to $\boldsymbol{A}$) j if j's equilibrium action is non-robust. Let $\mathcal{C}(\boldsymbol{T}, \boldsymbol{A}) \subseteq \mathcal{N}_n$ be the set of strongly connected components of $\boldsymbol{D}$ (see §1 for a definition). For any $G \subseteq \mathcal{N}_n$, define

$$\mathcal{S}(G) = G \cup \big\{ k \in \mathcal{N}_n \colon \max_{j \in G} A_{jk}(1 - \mathcal{R}_k^c) = 1 \big\}.$$

This adds to G the set of agents with robust actions that are connected to G.

Definition 1. $\mathcal{S}(C)$ is a *strategic neighborhood* if $C \in \mathcal{C}(\boldsymbol{T}, \boldsymbol{A})$.

It is not hard to show that the set of strategic neighborhoods coincides with the set of weakly connected components of $\boldsymbol{D}$ (defined in §1).

Example 2. In Figure 1, agents with robust (non-robust) actions are colored gray (white). Notice $\boldsymbol{D}$ has five strongly connected components, which are the "islands"

that result from deleting all links involving the gray agents: $\{1,2,4\},\{3\},\{5\},\{6\},\{7,8,9\}$. For example, $\{1,2,4\}$ is a strongly connected component, since we can travel from any agent to another through a path of agents with non-robust actions. On the other hand, $\{5,6\}$ is not a strongly connected component because $D_{65} = 0$. To obtain the strategic neighborhoods, we add to each component the gray agents connected to it, resulting in $\{1,\ldots,5\},\{3\},\{5\},\{5,6\},\{5,7,8,9\}$. Observe that these are the weakly connected components of $\boldsymbol{D}$. For example, $\{5,6\}$ is such a component because $D_{56} = 1$. Note that the strongly connected components partition $\mathcal{N}_n$, whereas the strategic neighborhoods do not.

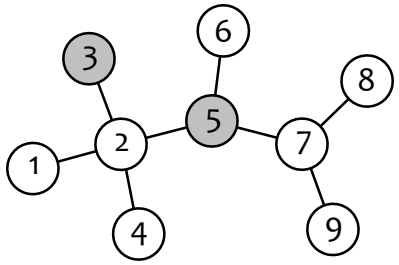

Figure 1: Gray agents have robust actions, white non-robust.

Recall that for any action profile $\boldsymbol{Y}$, Y_G is the subprofile $(Y_i)_{i\in G}$ for any $G \subseteq \mathcal{N}_n$. Our algorithm exploits the following property of strategic neighborhoods, which is a consequence of Assumption 1:

$$Y_{\mathcal{S}(C)} \in \mathcal{E}_{NE}(T_{\mathcal{S}(C)}, A_{\mathcal{S}(C)}) \quad \forall\, \boldsymbol{Y} \in \mathcal{E}_{NE}(\boldsymbol{T}, \boldsymbol{A}). \tag{2.4}$$

That is, for any Nash equilibrium $\boldsymbol{Y}$, the subprofile $Y_{\mathcal{S}(C)}$ is a Nash equilibrium in the game with only players in $\mathcal{S}(C)$. A formal proof is given in Lemma SA.2.2. For intuition, consider Figure 1. Because agent 1 has a non-robust action, its optimality may be affected by changes in the action of her neighbor, agent 2. Likewise, agent 2's action may be affected by changes in those of 3 and 5. However, 3 and 5 have robust actions and consequently maintain the same equilibrium action regardless of those of other agents, say 6 and 7. Therefore, if actions for agents in the strategic neighborhood $\{1,2,3,4,5\}$ are at equilibrium, then they remain optimal even after deleting agents $\{6,7,8,9\}$ from the network.

2.2 Algorithm

Our algorithm exploits property (2.4) to decompose $\mathcal{E}_{NE}(\boldsymbol{T}, \boldsymbol{A})$ into the Cartesian product of equilibrium sets on smaller subnetworks, namely those on strategic neighborhoods. For each $C \in \mathcal{C}(\boldsymbol{T}, \boldsymbol{A})$, we need to compute $\mathcal{E}_{NE}(T_{\mathcal{S}(C)}, A_{\mathcal{S}(C)})$. Under assumptions stated in the next subsection, we prove that it is feasible to compute these sets via exhaustive search because the size of the largest strategic neighborhood is $O_p(\log n)$. To combine these sets and obtain $\mathcal{E}_{NE}(\boldsymbol{T}, \boldsymbol{A})$, we have to account for the fact that strategic neighborhoods are not necessarily disjoint. Specifically, if $i \in \mathcal{S}(C) \cap \mathcal{S}(C')$ for two distinct components C, C', then we need to decide whether i's action should be dictated by profiles in $\mathcal{E}_{NE}(T_{\mathcal{S}(C)}, A_{\mathcal{S}(C)})$ or $\mathcal{E}_{NE}(T_{\mathcal{S}(C')}, A_{\mathcal{S}(C')})$. Fortunately, since C, C' must be disjoint by virtue of being strongly connected components, it follows that i's equilibrium action is necessarily robust and therefore the same across all equilibria in these two sets.

In order to succinctly state the algorithm, we need some additional notation. For $G \subseteq H \subseteq \mathcal{N}_n$, let

$$\mathcal{E}_{NE}(T_H, A_H)|_G = \big\{\boldsymbol{Y} \in \{0,1\}^{|G|} \colon \boldsymbol{Y} = Y'_G \text{ for some } \boldsymbol{Y}' \in \mathcal{E}_{NE}(T_H, A_H)\big\},$$

which simply drops from each equilibrium action profile in $\mathcal{E}_{NE}(T_H, A_H)$ the actions corresponding to agents in $H \backslash G$. Next, for any $C \in \mathcal{C}(\boldsymbol{T}, \boldsymbol{A})$ and $i \in \mathcal{S}(C)$, let

$$\pi(i; C) = \sum_{j \in \mathcal{S}(C)} \mathbf{1}\{j \leqslant i\}, \tag{2.5}$$

the number of agents in $\mathcal{S}(C)$ with label less than or equal to i.[4] Finally, define

$$\mathcal{Y}(\mathcal{S}(C), \boldsymbol{T}) = \big\{\boldsymbol{Y} \in \{0,1\}^{|\mathcal{S}(C)|} \colon Y_{\pi(i;C)} = 1 \text{ if } \inf_s U(s, T_i) > 0 \text{ and } Y_{\pi(i;C)} = 0 \text{ if } \sup_s U(s, T_i) \leqslant 0 \ \ \forall\, i \in \mathcal{S}(C)\big\}. \tag{2.6}$$

This is a subset of all possible action profiles for agents on $\mathcal{S}(C)$. Each profile in (2.6)

[4]This definition has the following purpose. Let $\boldsymbol{Y} \in \{0,1\}^{|\mathcal{S}(C)|}$ be an action profile for agents in $\mathcal{S}(C)$. Throughout the paper, our convention is that the components of $\boldsymbol{Y}$ are arranged in increasing order of the label of the corresponding agent. For example, if $\mathcal{S}(C) = \{1, 5\}$, then the first (second) component of $\boldsymbol{Y}$ dictates the action of agent 1 (5). Hence, under this convention, the action of agent $i \in \mathcal{S}(C)$ according to profile $\boldsymbol{Y}$ is given by $Y_{\pi(i;C)}$.

fixes the actions of agents with robust actions ($\mathcal{R}_i^c = 0$) at their dominant strategies (1 if $\inf_s U(s, T_i) > 0$ and 0 if $\sup_s U(s, T_i) \leqslant 0$). The actions of agents with non-robust actions are not fixed and vary freely across profiles in this set. We state our proposed procedure in Algorithm 1.

Algorithm 1: Procedure for computing the set of pure-strategy Nash equilibria.

Input: $\boldsymbol{T}$, $\boldsymbol{A}$, $U(\cdot)$
Output: $\mathcal{E}_{NE}(\boldsymbol{T}, \boldsymbol{A})$

1 *Compute* $\boldsymbol{D}$ *and then* $\{\mathcal{S}(C) \colon C \in \mathcal{C}(\boldsymbol{T}, \boldsymbol{A})\}$ *using depth-first search of* $\boldsymbol{D}$.
2 *Compute each* $\mathcal{E}_{NE}(T_{\mathcal{S}(C)}, A_{\mathcal{S}(C)})$ *using exhaustive search:*
for $C \in \mathcal{C}(\boldsymbol{T}, \boldsymbol{A})$ **do**
  $\mathcal{E}_{\mathcal{S}(C)} \leftarrow \varnothing$
  for $\boldsymbol{Y} \in \mathcal{Y}(\mathcal{S}(C), \boldsymbol{T})$ **do**
    if $Y_{\pi(k;C)} = \mathbf{1}\{U(S_k(\boldsymbol{Y}, T_{\mathcal{S}(C)}, A_{\mathcal{S}(C)}), T_k) > 0\}$ *for all* $k \in C$ **then**
      $\mathcal{E}_{\mathcal{S}(C)} \leftarrow \mathcal{E}_{\mathcal{S}(C)} \cup \{\boldsymbol{Y}\}$
    end
  end
  $\mathcal{E}_{NE}(T_{\mathcal{S}(C)}, A_{\mathcal{S}(C)}) \leftarrow \mathcal{E}_{\mathcal{S}(C)}$
end
3 *Combine equilibrium sets:*
if $\mathcal{E}_{NE}(T_{\mathcal{S}(C)}, A_{\mathcal{S}(C)}) \neq \varnothing \ \forall\, C \in \mathcal{C}(\boldsymbol{T}, \boldsymbol{A})$ **then**
  $\mathcal{E}_{NE}(\boldsymbol{T}, \boldsymbol{A}) \leftarrow (\times_{C \in \mathcal{C}(\boldsymbol{T}, \boldsymbol{A})} \mathcal{E}_{NE}(T_{\mathcal{S}(C)}, A_{\mathcal{S}(C)})|_C)$
else $\mathcal{E}_{NE}(\boldsymbol{T}, \boldsymbol{A}) \leftarrow \varnothing$
.

Remark 1 (Explanation of Algorithm 1)**.** Line 1 computes the set of strategic neighborhoods. As discussed in the previous subsection, this is the same as the set of weakly connected components of $\boldsymbol{D}$, which can be efficiently computed using well-known algorithms based on depth-first search. See for example the Matlab function `graphconncomp()` or the NetworkX Python function `weakly_connected_components()`. The computational complexity is $O(n+L)$ where L is the number of links in $\boldsymbol{D}$ (Kleinberg and Tardos, 2006). This assumes the graph is implemented using an adjacency list, which is the efficient format for sparse graphs, the focus of this paper (see Remark 4).

The for-loops in line 2 can be parallelized. Each iteration of the outer loop searches over strongly connected components of $\boldsymbol{D}$, or equivalently, strategic neighborhoods. For each such component C, the inner loop performs an exhaustive search over candi-

date action profiles in $\mathcal{Y}(\mathcal{S}(C), \boldsymbol{T})$ to find those satisfying the equilibrium conditions (2.1). Any other algorithm guaranteed to compute the equilibrium set may be used in place of exhaustive search (e.g. Daskalakis and Papadimitriou, 2006). In principle, one has to search through all possible profiles in $\{0,1\}^{|\mathcal{S}(C)|}$ to find the equilibrium set. However, it suffices to fix the actions of agents with robust actions ($\mathcal{R}_i^c = 0$) at their optimal values. This is precisely what we do by only searching through $\mathcal{Y}(\mathcal{S}(C), \boldsymbol{T})$ instead of all conceivable action profiles in $\{0,1\}^{|\mathcal{S}(C)|}$.

Line 3 shows how to assemble the equilibrium sets to obtain $\mathcal{E}_{NE}(\boldsymbol{T}, \boldsymbol{A})$. For clarity,

$$\bigtimes_{C \in \mathcal{C}(\boldsymbol{T}, \boldsymbol{A})} \mathcal{E}_{NE}(T_{\mathcal{S}(C)}, A_{\mathcal{S}(C)})\big|_C = \big\{\boldsymbol{Y} \in \{0,1\}^n \colon Y_C \in \mathcal{E}_{NE}(T_{\mathcal{S}(C)}, A_{\mathcal{S}(C)})|_C \ \forall\, C \in \mathcal{C}(\boldsymbol{T}, \boldsymbol{A})\big\}. \quad (2.7)$$

This is well-defined because $\mathcal{C}(\boldsymbol{T}, \boldsymbol{A})$ partitions $\mathcal{N}_n$.

Example 3. Consider Figure 1. As discussed in Example 2, there are five strategic neighborhoods that line 2 iterates over: $S_1 = \{1, \ldots, 5\}$, $S_2 = \{3\}$, $S_3 = \{5\}$, $S_4 = \{5, 6\}$, and $S_5 = \{5, 7, 8, 9\}$. The set of action profiles $\mathcal{Y}(S_1, \boldsymbol{T})$ contains 2^3 elements because agents 3 and 5 have robust actions. Suppose $(1, 1, 0, 1, 0)$ is the unique Nash equilibrium on S_1, where the actions are listed in increasing order of agents' labels. That is, $\mathcal{E}(T_{S_1}, A_{S_1}) = \{(1, 1, 0, 1, 0)\}$. Then by definition of robustness, we know that (0) is the unique Nash equilibrium on S_2 and S_3. Suppose $(0, 1, 0, 0)$ and $(0, 0, 1, 1)$ are the only Nash equilibria on S_5. Note that the first elements of these vectors must be 0, since agent 5's action is robust. Finally, note that there can only be a single equilibrium on S_4, since agent 6 simply best responds to agent 5; suppose that equilibrium is $(0, 1)$. Then

$$\mathcal{E}(\boldsymbol{T}, \boldsymbol{A}) = \{(1, 1, 0, 1, 0, 1, 1, 0, 0), (1, 1, 0, 1, 0, 1, 0, 1, 1)\}.$$

Remark 2 (Diagnostic for Computational Feasibility)**.** A very simple and quick way to assess the feasibility of Algorithm 1 is to compute $\Delta \equiv |C_1^*|$, where C_1^* is the largest strongly connected component of $\boldsymbol{D}$. This is a byproduct of line 1 of Algorithm 1. The significance of Δ is that the most computationally intensive step of Algorithm 1 is exhaustive search over $\mathcal{Y}(\mathcal{S}(C_1^*), \boldsymbol{T})$. The size of this set is 2^Δ, since all agents in C_1^*

have non-robust actions by definition. In the next subsection, we provide conditions under which $|\Delta| = O_p(\log n)$, in which case 2^Δ has size polynomial in n. As we will discuss, if these conditions are violated, Δ can instead be order n, in which case computing the equilibrium set is infeasible using our approach.[5]

Remark 3 (Strategic Complements)**.** As discussed in §1, under strategic complements, the set of equilibria has a lattice structure, and the extremal equilibria can be computed in polynomial time. However, finding all equilibria requires exhaustively searching all action profiles between these extremes. Our algorithm can be modified to speed this process up simply by restricting the search space $\mathcal{Y}(\mathcal{S}(C), \boldsymbol{T})$ in line 2 to profiles between these extremes.

2.3 Assumptions

We state two conditions required by our main result in the next subsection, which provides bounds on the complexity of Algorithm 1. It should be emphasized that, in practice, all that matters for computational feasibility is the size of the largest component of $\boldsymbol{D}$, as discussed in Remark 2. The purpose of these conditions is to give a theoretical sense of what is required of the data-generating process for this number to be typically small in practice.

The key condition for our result is a restriction on the strength of strategic interactions. Leung (2019) uses this condition to prove a CLT for graphical games. An analogous condition is used by Leung and Moon (2019) to obtain a CLT for network formation games. The main theoretical contribution of this paper is to show that this condition also enables feasible computation of $\mathcal{E}_{NE}(\boldsymbol{T}, \boldsymbol{A})$.

To motivate the condition, consider the standard linear-in-means model of social interactions (Bramoullé et al., 2009)

$$Y_i = \alpha + \sum_j \tilde{A}_{ij} Y_j \beta + \sum_j \tilde{A}_{ij} X_j' \delta + X_i' \gamma + \varepsilon_i,$$

where $\tilde{\boldsymbol{A}}$ is the row-normalized version of the adjacency matrix $\boldsymbol{A}$, obtained by di-

[5]Exhaustive search on the largest component will typically be the only computationally intensive step of the algorithm. What might be considered a "folk theorem" in random graph theory is that the size of the second largest component is typically much smaller than that of the first. For example, in the Erdős-Rényi model, its size is $O_p(\log n)$ (Bollobás, 2001, Ch. 6). Hence, exhaustive search is typically trivial on all components but the largest.

viding each row by its row sum. This can be microfounded as the best-response function of a linear-quadratic model (Calvó-Armengol et al., 2009). Existence of an equilibrium requires $|\beta\lambda_{\max}(\tilde{\boldsymbol{A}})| < 1$, where $\lambda_{\max}(\tilde{\boldsymbol{A}})$ is the largest eigenvalue of the adjacency matrix. Under row-normalization, this simplifies to

$$|\beta| < 1, \tag{2.8}$$

which restricts the strength of the endogenous peer effect. Without this condition, best responses would be "explosive" in that an agent would respond to the actions of her neighborhoods with a larger action in expectation, and her neighbors would respond with still larger actions, etc. Therefore, no equilibrium could exist.

The assumption we impose is an analog of (2.8) for models with discrete outcomes. The statement of the assumption is more complicated, but this is unavoidable due to the nonlinearity and generality of the model. Recall the non-robustness indicator $\mathcal{R}_i^c$ from §2.1, and notice that

$$\mathbf{E}[\mathcal{R}_i^c] = \mathbf{P}\big(\sup_s U(s, T_i) > 0\big) - \mathbf{P}\big(\inf_s U(s, T_i) > 0\big) \tag{2.9}$$

(assuming measurability). This corresponds to the partial-equilibrium marginal effect of changing statistics $S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A})$ from their minimizing to their maximizing value on i's propensity to choose action 1. Hence, it measures the strength of strategic interactions, analogous to (2.8). For instance, in Example 1, if $\beta > 0$, then (2.9) $= \mathbf{P}(0 \leqslant \xi(T_i) < \beta)$, which is clearly monotonic in β.

Recall that $\tau_i \in \mathbb{R}^d$ is the subvector of T_i relevant for network formation (2.2). Let μ be a measure on $\mathbb{R}^d$ (which need not integrate to 1). For a given $k \leqslant d$, consider a partition of the type space $\mathbb{R}^d = \mathbb{R}^k \times \mathbb{R}^{d-k}$, and let μ_k and μ_{-k} be the marginals of μ over $\mathbb{R}^k$ and $\mathbb{R}^{d-k}$, respectively. For any $h\colon \mathbb{R}^d \to [0, 1]$, define the mixed norm

$$\|h\|_{\mathbf{m},k} = \sup_{t_k \in \mathbb{R}^k} \left(\int_{\mathbb{R}^{d-k}} h(t_k, t_{-k})^2 \, \mathrm{d}\mu_{-k}(t_{-k}) \right)^{1/2}.$$

Assumption 2 (Strength of Interactions)**.** *There exist a measure μ on $\mathbb{R}^d$ and function $\varphi\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ such that, for some integer $0 \leqslant k \leqslant d$, (a) μ is a product*

measure on $\mathbb{R}^k \times \mathbb{R}^{d-k}$*, (b)*

$$n\mathbf{P}(D_{ij} = 1 \mid \tau_i = t) \leqslant \int_{\mathbb{R}^d} \varphi(t, t')\, d\mu(t')$$

for all $t \in \mathbb{R}^d$ *and* $n \in \mathbb{N}$*, and (c)*

$$\|\lambda\|_{\mathbf{m},k} < 1 \quad \textit{for} \quad \lambda(t) \equiv \int_{\mathbb{R}^k} \left(\int_{\mathbb{R}^{d-k}} \varphi(t, (t'_k, t'_{-k}))^2 \, d\mu_{-k}(t'_{-k}) \right)^{1/2} d\mu_k(t'_k).$$

The statement of this assumption is admittedly complicated, but this is in order to allow for a broad class of network formation models. By Jensen's inequality, (b) and (c) imply that $n^{-1} \sum_i \sum_{j \neq i} \mathbf{E}[D_{ij}]$, the expected *degree* of $\boldsymbol{D}$, is less than one in the limit. If this is violated, the size of the largest component of $\boldsymbol{D}$ can be large, and computation is infeasible. In practice, as discussed in Remark 2, a direct way of assessing computational feasibility is just to compute Δ defined in that remark.

The next example illustrates how Assumption 2 is fully analogous to (2.8).

Example 4. Consider the simple case of an exogenous network, where $A_{ij} \perp\!\!\!\perp \mathcal{R}_j^c$.[6] Since types are independent,

$$n\mathbf{P}(D_{ij} = 1 \mid \tau_i = t) = \mathbf{E}[\mathcal{R}_j^c]\, n\mathbf{P}(A_{ij} = 1 \mid \tau_i = t). \tag{2.10}$$

Consider the following inhomogeneous random graph model for $\boldsymbol{A}$:

$$\mathbf{P}(A_{ij} = 1 \mid \tau_i, \tau_j) = \rho_n h(\tau_i, \tau_j)$$

for some bounded function h and $\rho_n = \kappa/n$ for some $\kappa > 0$.[7] We have

$$n\mathbf{P}(A_{ij} = 1 \mid \tau_i = t) = \int_{\mathbb{R}^d} \kappa\, h(t, t')\, \mathrm{d}\mu^*(t'), \tag{2.11}$$

for all n and t, where μ^* is the distribution of τ_1.

To put this in the setup of Assumption 2, choose $k = 0$, $\mu_{-k} = \mu^*$. Then for

[6] Exogeneity holds if T_i can be partitioned into two independent subvectors, τ_i and X_i, where $U(s, T_i)$ only depends on T_i through X_i, while $A_{ij} = g_n(\tau_i, \tau_j, \zeta_{ij})$, with $\zeta_{ij} \perp\!\!\!\perp T_j$.

[7] This model has been extensively studied by probabilists (Bollobás et al., 2007). A large literature on community detection in statistics uses the inhomogeneous random graph (also called the stochastic block model) as the data-generating process (Bickel et al., 2011).

$\varphi(t,t') = \kappa \mathbf{E}[\mathcal{R}_j^c] h(t,t')$,

$$\|\lambda\|_{\mathbf{m},k} = \mathbf{E}[\mathcal{R}_j^c] \, \|\kappa h(\tau_1,\tau_2)\|_2,$$

where $\|\cdot\|_2$ is the L_2-norm. Therefore, Assumption 2 requires

$$\mathbf{E}[\mathcal{R}_j^c] < \left(\|\kappa h(\tau_1,\tau_2)\|_2\right)^{-1}. \tag{2.12}$$

Recall from (2.9) that the left-hand side measures the strength of strategic interactions. By Jensen's inequality, $\|\kappa h(\tau_1,\tau_2)\|_2 \geqslant \lim_{n\to\infty} n\mathbf{E}[A_{ij}]$, the latter being the limiting expected degree of any agent. Hence, (2.12) implies that *the strength of strategic interactions is bounded above the inverse of the limiting expected degree*:

$$\mathbf{E}[\mathcal{R}_j^c] < \left(\lim_{n\to\infty} \mathbf{E}\left[\sum_j A_{ij}\right]\right)^{-1}.$$

To see that this is analogous to (2.8), consider a model that does not row-normalize the adjacency matrix, meaning we replace $\tilde{\boldsymbol{A}}$ with $\boldsymbol{A}$. Since $\lambda_{\max}(\boldsymbol{A}) \leqslant \max_i \sum_j A_{ij}$, a sufficient condition for $|\beta\lambda_{\max}(\boldsymbol{A})| < 1$ is

$$|\beta| < \left(\max_i \sum_j A_{ij}\right)^{-1}, \tag{2.13}$$

which is directly analogous to the prior equation. Furthermore, Assumption 2 and (2.13), as well as (2.8) in the row-normalized model, have the same behavioral implication that an exogenous one-unit change in the average agent's outcome causes the total outcome of her neighbors to change by less than one unit. Remark SA.1.1 in §SA.1 elaborates on this point. Finally, note that this example sets $k = 0$; for a model in which $k > 0$ and μ_k is nontrivial, see Example SA.1.1 in §SA.1.

Remark 4 (Sparsity)**.** In Example 4, a necessary condition for (2.12) is that either $\mathbf{E}[\mathcal{R}_j^c] = 0$, in which case there are no strategic interactions, or $\|\kappa h(\tau_1,\tau_2)\|_2 < \infty$, which, from the discussion in Example 4, means that the limiting expected degree of any agent is finite. This is a common notion of network *sparsity* for $\boldsymbol{A}$ (Barabási, 2015). It formalizes the well-known stylized fact that, for most real-world social networks, the number of connections involving the typical agent is significantly smaller

than the network size (Chandrasekhar, 2016).

The next assumption imposes mild regularity conditions.

Assumption 3 (Regularity)**.**

(a) $\sup_{t,t'} \mathbf{P}(D_{ij} = 1 \mid \tau_i = t, \tau_j = t') < 1$ *for any* n*.*

(b) $\sup_t \int_{\mathbb{R}^k} \left(\int_{\mathbb{R}^{d-k}} \varphi(t, (t'_k, t'_{-k}))^2 \, d\mu_{-k}(t'_{-k}) \right)^{1/2} d\mu_k(t'_k) < \infty$*.*

(c) Either $\varphi(t, t') = 0$ *for any* $t, t' \in \mathbb{R}^d$*, or*

$$\inf_{t \in \mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(t, t') \, d\mu(t') > 0$$

for $\varphi(\cdot)$ *and* μ *defined in Assumption 2.*

In Example 4 both (a) and (b) immediately hold because $h(\cdot)$ is bounded. In part (c), if $\varphi(t, t') = 0$ everywhere, then this corresponds to a model either with an empty network ($A_{ij} = 0$ a.s.) or with no strategic interactions ($\mathcal{R}_j^c = 0$ a.s.). For models with strategic interactions, part (c) implies that the network is asymptotically nondegenerate. For instance, in Example 4

$$\inf_{t \in \mathbb{R}^d} \int_{\mathbb{R}^d} \varphi(t, t') \, \mathrm{d}\mu(t') = \mathbf{E}[\mathcal{R}_j^c] \inf_{t \in \mathbb{R}^d} \int_{\mathbb{R}^d} \kappa h(t, t') \, \mathrm{d}\mu(t').$$

The infimum term on the right-hand side is the infimum over t of the conditional limiting expected degree of an agent of type t. If this is strictly positive, it means the expected number of connections involving any agent is nonzero in the large-network limit.

2.4 Algorithmic Complexity

We next state our main result, the proof of which is given in §SA.2.

Theorem 1. *Suppose evaluating* $\mathcal{R}_i^c$ *and* $U_i(S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i)$ *have the same complexity for any* i*. Under Assumptions 1–3, Algorithm 1 computes* $\mathcal{E}_{NE}(\boldsymbol{T}, \boldsymbol{A})$ *in* $O_p(n^{1+q})$ *evaluations of the payoff function for* $q > \log 2 / \log \|\lambda\|_{\mathbf{m},k}^{-1}$*, where* $\|\lambda\|_{\mathbf{m},k}$ *is defined in Assumption 2.*

Remark 5. The assumption that $\mathcal{R}_i^c$ and $U_i(S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i)$ have the same complexity is a mild requirement, since, in practice, the payoff function is typically monotonic in each component of $S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A})$. If the assumption fails, then the complexity is instead of order $n^{1+q} + \sum_{i=1}^n \mathcal{R}_i^c$.

Remark 6. The theorem says that Algorithm 1 has polynomial complexity in models for which $\|\lambda\|_{\mathbf{m},k}$ lies below the "unit root" of 1. How far this parameter lies below the unit root determines the speed of computation through the order of the polynomial q. If $\|\lambda\|_{\mathbf{m},k} = 0.5$, then q can be chosen arbitrarily close to one. If instead it is closer to one, then the computational speed can be slow. This is because the size of the largest component of $\boldsymbol{D}$ increases with the magnitude of this parameter, as shown in simulations (§3); Remark 8 below provides some intuition for this fact. As discussed in Remark 2, computing the equilibrium set on the strategic neighborhood corresponding to the largest component is the main computational bottleneck.

Remark 7. The key step of the proof is establishing a polynomial bound on the complexity of line 2 of the algorithm. This results from an exponential tail bound on the size of any arbitrary component of the network $\boldsymbol{D}$. The proof technique is to traverse the component using a breadth-first search, where we start at an arbitrary agent, branch to her neighbors in $\boldsymbol{D}$, then branch to their neighbors, etc. The number of agents traversed at each step can be stochastically bounded by a multi-type Galton-Walton branching process. Then using an argument due to Turova (2012), we derive exponential bounds on the size of this process. This line of argument is commonly used in random graph theory to study the phase transition of the giant component (e.g. Janson et al., 2011, Ch. 3).

Remark 8. The key condition is Assumption 2. To understand its role, imagine "$n = \infty$." From the discussion in the previous subsection, we can roughly interpret $\|\lambda\|_{\mathbf{m},k}$ as the expected degree of an arbitrary agent in $\boldsymbol{D}$ (formally it is an upper bound by Jensen's inequality). For the branching process in Remark 7, as we branch from neighborhood to neighborhood, the number of additional agents traversed progressively shrinks if the expected degree is less than one because the process is below its "replacement rate." In this case, the branching process eventually reaches extinc-

tion.[8] Hence, its total size is stochastically bounded, and tail bounds may be obtained. If the expected degree instead exceeds one, then the expected size of the branching process can grow exponentially without bound because each agent is replaced with more than one neighbor in expectation. This case corresponds to having a "giant component" in the network, meaning the size of the largest component is order n.[9] Then the number of candidate action profiles to search through is exponential in n, as discussed in Remark 2, so exhaustive search is infeasible.

Remark 9. Theorem SA.4.2 in the supplemental appendix generalizes Theorem 1, allowing $\boldsymbol{A}$ to be drawn from a stochastic model of network formation with strategic interactions.

2.5 Multinomial Choice

Our previous results pertain to games with binary action spaces. We next consider the extension to $K+1>2$ unordered actions, which we arbitrarily label $\{0,\dots,K\}$. Let $U_k(S_i(\boldsymbol{Y},\boldsymbol{T},\boldsymbol{A}),T_i)$ be i's utility from choosing action k, where

$$S_i(\boldsymbol{Y},\boldsymbol{T},\boldsymbol{A}) \equiv S(Y_{-i},T_i,T_{-i},A_i,A_{-i}), \tag{2.14}$$

and $S(\cdot)$ is a function with range $\mathbb{R}^{d_s}$ satisfying Assumption 1. An action profile $\boldsymbol{Y}=(Y_i)_{i=1}^n$ is a pure-strategy Nash equilibrium if for every $i\in\mathcal{N}_n$,

$$Y_i=k \quad \text{if and only if} \quad U_k(S_i(\boldsymbol{Y},\boldsymbol{T},\boldsymbol{A}),T_i) > U_\ell(S_i(\boldsymbol{Y},\boldsymbol{T},\boldsymbol{A}),T_i) \quad \forall\,\ell\neq k. \tag{2.15}$$

Example 5. Suppose $T_i=((X_{ik}',\varepsilon_{ik}))_{k=0}^K$, where X_{ik} is observed and ε_{ik} unobserved. Let

$$U_k(S_i(\boldsymbol{Y},\boldsymbol{T},\boldsymbol{A}),T_i) = \theta_1 + X_{ik}'\theta_2 + \sum_\ell \beta_{\ell,k}\frac{\sum_j A_{ij}\mathbf{1}\{Y_j=\ell\}}{\sum_j A_{ij}} + \varepsilon_{ik}$$

This allows payoffs from action k to depend on the fraction of neighbors choosing action ℓ for any ℓ, and the peer effect $\beta_{\ell,k}$ can vary across ℓ and k.

[8] In a single-type branching process, the expected total population size is $\sum_{t=0}^\infty \mu^k$, where μ is the expected number of offspring produced by any particle (Mode, 1971). Hence, when $\mu<1$, this is $(1-\mu)^{-1}<\infty$. The parameter μ is analogous to $\|\lambda\|_{\mathbf{m},k}$ in our context.

[9] For a formal result for the case of inhomogeneous random graphs, see Theorem 3.1 of Bollobás et al. (2007).

Algorithm 1 and Theorem 1 can be extended to this setting if we redefine the non-robustness indicator (2.3) as

$$\mathcal{R}_i^c = \mathbf{1}\big\{ \inf_s \min_{\ell \neq k} (U_k(s, T_i) - U_\ell(s, T_i)) \leqslant 0 \quad \forall\, k \in \{0, \ldots, K\} \big\}. \tag{2.16}$$

To understand this, note that if $\mathcal{R}_i^c = 0$, then there is some action k such that the marginal utility of choosing k over any other ℓ is positive in any Nash equilibrium ($\inf_s \min_{\ell \neq k}(U_k(s, T_i) - U_\ell(s, T_i)) > 0$), in which case choosing k is the dominant strategy. This definition reduces to (2.3) in the binary choice setting where $K = 1$ and the payoff of action 0 is normalized to 0.

Algorithm 1 can be applied to compute the set of Nash equilibria under multinomial choice by redefining $\mathcal{R}_i^c$ as (2.16), the definition of equilibrium in line 2 of Algorithm 1 as (2.15), and

$$\begin{aligned} \mathcal{Y}(\mathcal{S}(C), \boldsymbol{T}) = \big\{ \boldsymbol{Y} \in \{0, \ldots, K\}^{|\mathcal{S}(C)|} \colon Y_{\pi(i;C)} = k \text{ if } \\ \inf_s \min_{\ell \neq k} (U_k(s, T_i) - U_\ell(s, T_i)) > 0,\ \forall\, i \in \mathcal{S}(C), k \in \{0, \ldots, K\} \big\}. \end{aligned}$$

Analogously to (2.6), this is the set of action profiles on $\mathcal{S}(C)$ that fix the actions of agents with robust actions at their dominant strategies. The resulting algorithm computes in $O_p(n^{1+q})$ time by Theorem 1, whose proof applies almost verbatim. Example SA.1.2 in §SA.1 illustrates the interpretation of Assumption 2 in the multinomial choice setting.

2.6 Ordered Choice

We next consider the case in which the action space $\{0, \ldots, K\}$ is ordered in the natural way. Let $U(y, S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i)$ denote i's utility from choosing action y, where $S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A})$ is defined as in (2.14). An action profile $\boldsymbol{Y} = (Y_i)_{i=1}^n$ is a pure-strategy Nash equilibrium if, for every $i \in \mathcal{N}_n$,

$$Y_i = \operatorname*{argmax}_{y \in \{0, \ldots, K\}} U(y, S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i). \tag{2.17}$$

Define $U(K+1, S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i) = U(-1, S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i) = -\infty$. Then by Lemma

1 of Aradillas-Lopez and Rosen (2019), (2.17) is equivalent to

$$U(Y_i, S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i) \\ \geqslant \max\{U(Y_i+1, S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i), U(Y_i-1, S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i)\} \quad (2.18)$$

if $U(\cdot)$ satisfies a strict concavity condition in its first argument (their Restriction SR(i)). Concavity is a mild restriction guaranteeing each agent i has an a.s. unique best response to any profile of actions Y_{-i}. We maintain this assumption in what follows.

Example 6. Card and Giuliano (2013) estimate an empirical model of peer effects in risky behavior among teens. They consider an ordered response model with strategic interactions with $K = 2$, where 0 indicates no sexual activity, 1 intimate contact without intercourse, and 2 intercourse. Their model is equivalent to normalizing the payoff of action 0 to 0 and setting $U(1, S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i) = X_i'\beta + \varepsilon_i - c_1(Y_{-i}, \boldsymbol{A})$ and $U(2, S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i) = 2(X_i'\beta + \varepsilon_i) - c_1(Y_{-i}, \boldsymbol{A}) - c_2(Y_{-i}, \boldsymbol{A})$. Here $T_i = (X_i, \varepsilon_i)$, and $S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}) = (c_1(Y_{-i}, \boldsymbol{A}), c_2(Y_{-i}, \boldsymbol{A}))$ are cutoffs that depend on the actions of other agents. Then defining $U_i^* = X_i'\beta + \varepsilon_i$, a Nash equilibrium $\boldsymbol{Y}$ satisfies, for all agents i,

$$Y_i = \begin{cases} 0 & \text{if } U_i^* \leqslant c_1(Y_{-i}, \boldsymbol{A}), \\ 1 & \text{if } c_1(Y_{-i}, \boldsymbol{A}) < U_i^* \leqslant c_2(Y_{-i}, \boldsymbol{A}), \\ 2 & \text{if } c_2(Y_{-i}, \boldsymbol{A}) < U_i^*. \end{cases}$$

Examples of cutoffs are

$$c_1(Y_{-i}, \boldsymbol{A}) = \alpha_1 - \gamma_1 \frac{\sum_j A_{ij} \mathbf{1}\{Y_j \geqslant 1\}}{\sum_j A_{ij}}, \quad c_2(Y_{-i}, \boldsymbol{A}) = \alpha_2 - \gamma_2 \frac{\sum_j A_{ij} \mathbf{1}\{Y_j = 2\}}{\sum_j A_{ij}},$$

where $\alpha_2 > \alpha_1$ and $\alpha_1 < \alpha_2 - \gamma_2$ to ensure strict concavity. This example coincides with the baseline specification of Card and Giuliano (2013) in their setting where a network consists of only two linked agents. Under strategic complements ($\gamma_1, \gamma_2 \geqslant 0$), it says that agents choose riskier behaviors the higher the share of their friends choosing riskier behaviors.

We extend Algorithm 1 and Theorem 1 to ordered choice by redefining (2.3) as

$$\mathcal{R}_i^c = \mathbf{1}\Big\{ \inf_s \big(U(y,s,T_i) - \max\big\{U(y+1,s,T_i), U(y-1,s,T_i)\big\}\big) < 0 \\ \forall y \in \{0,\dots,K\}\Big\}.$$

Similar to (2.16), if $\mathcal{R}_i^c = 0$, then there is some action y such that the marginal utility of choosing y over any other action is positive in any Nash equilibrium, so that choosing y is the dominant strategy. In Example 6,

$$\mathcal{R}_i^c = \mathbf{1}\big\{X_i'\beta + \varepsilon_i - \alpha_1 + \max\{\gamma_1, 0\} > 0\big\}\mathbf{1}\big\{X_i'\beta + \varepsilon_i - \alpha_2 + \min\{\gamma_2, 0\} < 0\big\} \\ \times \mathbf{1}\big\{\min\{-(X_i'\beta + \varepsilon_i) + \alpha_2 - \max\{\gamma_2, 0\}, X_i'\beta + \varepsilon_i - \alpha_1 + \min\{\gamma_1, 0\}\} < 0\big\}.$$

Algorithm 1 can be applied to compute the set of Nash equilibria by redefining $\mathcal{R}_i^c$ as above, the definition of equilibrium in line 2 of Algorithm 1 as (2.18), and

$$\mathcal{Y}(\mathcal{S}(C), \boldsymbol{T}) = \big\{\boldsymbol{Y} \in \{0,\dots,K\}^{|\mathcal{S}(C)|} \colon Y_{\pi(i;C)} = y \text{ if} \\ \inf_s \big(U(y,s,T_i) - \max\big\{U(y+1,s,T_i), U(y-1,s,T_i)\big\}\big) \geqslant 0, \\ \forall\, i \in \mathcal{S}(C), y \in \{0,\dots,K\}\big\}.$$

Analogously to (2.6), this is the set of action profiles on $\mathcal{S}(C)$ where we fix the actions of agents with robust actions at their optimal choices. Theorem 1, whose proof applies almost verbatim to the resulting algorithm, implies its computational complexity is $O_p(n^{1+q})$.

3 Numerical Illustrations

This section illustrates the performance of our algorithms on graphical games estimated in Card and Giuliano (2013) and Xu (2018). We also show empirically the relationship discussed in Remark 8 between computational feasibility and the parameter $\|\lambda\|_{\mathbf{m},k}$ in Assumption 2 that controls the strength of strategic interactions.

3.1 Binary Choice

We consider a binary graphical game inspired by Xu (2018), who estimates a model of college attendance with social interactions. He specifies net payoffs from attendance as

$$U(S_i(\boldsymbol{Y}, \boldsymbol{T}, \boldsymbol{A}), T_i) = X_i'\theta + \beta \frac{\sum_j A_{ij} Y_j}{\sum_j A_{ij}} + \varepsilon_i,$$

which corresponds to Example 1 with a linear model for the threshold. Here Y_i is student i's college attendance decision, and A_{ij} represents friendship between i and j. We apply Algorithm 1 to compute the set of pure-strategy Nash equilibria using his estimated parameters.

Xu (2018) uses data from the restricted-use sample of Add Health (Harris and Udry, 2018), only using the three largest schools (831 student observations). We instead use all schools from the public-use sample, excluding students with missing data (1952 students). We construct the covariates X_i used in his application, and following his setup, draw ε_i from the logistic distribution.

The public-use sample contains information on degrees (number of friends of each student) but not the full network $\boldsymbol{A}$. We opt to simulate a network from a configuration model (Jackson, 2010, Ch. 4.1.4) calibrated to the empirical degree sequence of the Add Health network. This model approximately draws a network uniformly at random from the set of all networks such that the degree sequence matches the empirical degree sequence.[10]

For (θ, β), we take the estimate in Table 5, column AMLE(4) of Xu (2018) and add 0.2 to the estimated value of β. As noted in Xu's paper, the partial-equilibrium peer effects suggested by these estimates are fairly large and similar to those of other studies. With Xu's original estimates, for an agent with covariates equal to the mean covariate vector, the probability of college attendance increases by 11.83 percent if the fraction of friends attending increases from 0 to 0.5 in partial equilibrium. Using our sample and modified parameters, this marginal effect is instead 46 percent.

We compute the set of Nash equilibria for 100 simulation draws. Each simulation redraws random utility shocks and a network from the configuration model, while keeping covariates fixed. Shocks are independent across agents and of the network, as

[10] We use out-degrees (number of friends named by the ego) as degrees. We simulate a single network $\boldsymbol{A}$ on all observations. In reality, the data consists of several schools with few cross-school links.

in Xu's application. Computation is carried out on a laptop with a 2.6 GHz processor and 8 GB of memory. The algorithm is coded in Python and does not parallelize.

Table 1 reports the results. The rows give the mean, standard deviation, minimum, and maximum of the statistic across simulation draws. Note that the standard deviation for the average degree of $\boldsymbol{A}$ is almost zero because the degree distribution is given by the data and essentially fixed across draws. We see that the average outcome is similar across equilibria, and the number of equilibria is small, on average around three. The computation time is 1 second on average, although occasionally the largest component can have a larger number of agents, resulting in a longer computation time. In this case, the largest size is 19, and computing the corresponding set of equilibria takes about half a minute.

Table 1: Results

	$\bar{Y}$ Lower	$\bar{Y}$ Upper	# NE	Time	Δ	$\boldsymbol{D}$ Deg	$\boldsymbol{A}$ Giant	$\boldsymbol{A}$ Deg
Mean	0.141	0.142	3.2	1.1	8.4	0.488	1717.6	4.914
SD	0.008	0.008	3.0	5.3	3.4	0.040	1.440	0.004
Min	0.118	0.118	1.0	0.1	4.0	0.395	1715.0	4.902
Max	0.166	0.167	16.0	37.8	19.0	0.607	1719.0	4.921

$n = 1952$, 100 simulations. Column 1 is the smallest average outcome across simulations, column 2 the largest. Column 3 is the number of equilibria. Column 4 is computation time in seconds. Column 5 (7) is the size of the largest component of $\boldsymbol{D}$ ($\boldsymbol{A}$) and column 6 (8) the average degree of $\boldsymbol{D}$ ($\boldsymbol{A}$).

We next numerically illustrate the relationship between β and the size Δ of the largest component of $\boldsymbol{D}$. The motivation is that β controls the strength of strategic interactions $\|\lambda\|_{\mathbf{m},k}$. Our theoretical results require this parameter to be less than one, which ensures that the size of the largest component Δ grows only logarithmically with n. Recall that, in line 2 of Algorithm 1, the most computationally intensive step is computing the equilibrium set of the strategic neighborhood corresponding to the largest component of $\mathcal{C}(\boldsymbol{T}, \boldsymbol{A})$. As discussed in Remark 2, this requires searching over 2^{Δ} elements, so Δ determines the computational feasibility of our proposed algorithm. If β is too large, then Δ will be too large for feasible computation.

We simulate $\boldsymbol{D}$ according to the same model above, except with β increased by 0.6 rather than 0.2. As predicted by the theory, with a larger β (and hence larger $\|\lambda\|_{\mathbf{m},k}$), the average size of the largest component increases to 32. This corresponds

to a significant increase in 2^{Δ} and therefore the expected computation time. When we increase β by 1 rather than 0.6, the average size of the largest component balloons to 159, so computing the equilibrium set is infeasible in practice.

3.2 Ordered Choice

We repeat the exercise in the previous subsection, now using the ordered choice model of Card and Giuliano (2013) in Example 6. We construct their covariates X_i using the public-use sample of Add Health, whereas the authors utilize the restricted-use sample. After discarding observations with missing data, our sample size is 1184. We calibrate the values of $\beta, \gamma_1, \gamma_2$ using the third columns of Tables 3 and A3 of their paper. This corresponds to their specification allowing for peer effects but assuming independent random-utility shocks across agents. The authors do not report estimates of the intercepts α_1, α_2, so we set them to $-1.5, 1.5$, respectively.

For these parameter values and our data, given an agent with covariates equal to the mean covariate vector, their probability of engaging in no sexual activity (action 0) decreases by 23 percent if the fraction of friends engaging in at least some intimate contact (actions 1 or 2) increases from 0 to 0.5 in partial equilibrium. The corresponding partial-equilibrium effect on the probability of engaging in intercourse (action 2) is a 20 percent increase.

We compute the set of Nash equilibria for 100 simulation draws using modifications to Algorithm 1 discussed in §2.6. We draw random utility shocks ε_i from a standard normal distribution, following Card and Giuliano (2013). Independently of the shocks, we draw the network from a configuration model calibrated to the empirical degree sequence, as in the previous binary choice illustration. This setting differs from that of Card and Giuliano (2013) because we use friendship network data and assume our whole sample (1184 observations) is part of the same game, whereas the authors assume that each game consists of only two students who are best friends. Finally, due to the larger action space, we parallelize line 2 of the algorithm across 16 cores (8 GB memory per core, 2-3 GHz CPUs) for components C of size at least 12.

Table 2 reports the results of this exercise. The statistics for Δ are largely similar to the binary choice exercise. The computation time is 2 seconds on average, and the longest computation time is about 6 minutes, which corresponds to a simulation draw for which the size of the giant component of $\boldsymbol{D}$ is 16.

Table 2: Results

	$\bar{Y}$ Lower	$\bar{Y}$ Upper	# NE	Time	Δ	$\boldsymbol{D}$ Deg	$\boldsymbol{A}$ Giant	$\boldsymbol{A}$ Deg
Mean	1.224	1.226	2.1	6.2	6.3	0.401	985.5	4.515
SD	0.014	0.014	1.8	37.6	2.7	0.045	1.6	0.000
Min	1.188	1.188	1.0	0.1	2.0	0.230	981.0	4.515
Max	1.265	1.267	16.0	374.0	16.0	0.503	987.0	4.515

$n = 1184$, 100 simulations. Column 1 is the smallest average outcome across simulations, column 2 the largest. Column 3 is the number of equilibria. Column 4 is computation time in seconds. Column 5 (7) is the size of the largest component of $\boldsymbol{D}$ ($\boldsymbol{A}$) and column 6 (8) the average degree of $\boldsymbol{D}$ ($\boldsymbol{A}$).

4 Conclusion

We propose new algorithms for computing the set of Nash equilibria of graphical games. For models satisfying a restriction on the strength of strategic interactions, we show the algorithms typically complete in polynomial time, or more formally, in $O_p(n^c)$ evaluations of the payoff function for some $c > 1$. Our theory and simulation results suggest that the closer the restriction is to being violated, the larger the exponent c, and hence the more difficult it is to compute the equilibrium set.

The algorithms proceed by constructing small subgames for which it is feasible to compute the equilibrium set and then combining these sets. These subgames correspond to the components of a certain network, which can be quickly computed using well-known algorithms based on depth-first search. Under our conditions, we show that the largest component is typically small, scaling logarithmically with the network size. It then becomes feasible to exhaustively search the space of action profiles component by component to compute all equilibria.

In the supplemental appendix, we provide algorithms for computing equilibria of network formation games, including pairwise stable and Nash stable equilibria. The former solution concept only allows an agent to unilaterally deviate by deleting a single link, while the latter allows for unilateral deletion of multiple links. It is possible to consider refinements in the undirected setting, such as pairwise Nash stability, by combining aspects of both algorithms. We discuss how the algorithms can be applied to compute previously intractable sharp bounds on structural parameters based on moment inequalities in Sheng (2016).

References

Aradillas-Lopez, A. and A. Rosen, "Inference in Ordered Response Games with Complete Information," *CeMMaP working paper*, 2019.

Bajari, P., H. Hong, and S. Ryan, "Identification and Estimation of a Discrete Game of Complete Information," *Econometrica*, 2010, *78* (5), 1529–1568.

_ , _ , J. Krainer, and D. Nekipelov, "Computing Equilibria in Static Games of Incomplete Information using the All-Solution Homotopy," *Operations Research*, 2010, *58*, 237–45.

Barabási, A., *Network Science*, Cambridge University Press, 2015.

Bhattacharya, D., Pascaline D., and S. Kanaya, "Demand and Welfare Analysis in Discrete Choice Models with Social Interactions," *working paper*, 2019.

Bickel, P., A. Chen, and E. Levina, "The Method of Moments and Degree Distributions for Network Models," *Annals of Statistics*, 2011, *39* (5), 2280–2301.

_ and _ , "A Nonparametric View of Network Models and Newman–Girvan and other Modularities," *Proceedings of the National Academy of Sciences*, 2009, *106* (50), 21068–21073.

Bollobás, B., *Random Graphs* number 73, Cambridge University Press, 2001.

_ and O. Riordan, "Random Graphs and Branching Processes," in "Handbook of Large-Scale Random Networks," Springer, 2008, pp. 15–115.

Bollobás, B., S. Janson, and O. Riordan, "The Phase Transition in Inhomogeneous Random Graphs," *Random Structures and Algorithms*, 2007, *31* (1), 3–122.

Bramoullé, Y., H. Djebbari, and B. Fortin, "Identification of Peer Effects Through Social Networks," *Journal of Econometrics*, 2009, *150* (1), 41–55.

Brock, W. and S. Durlauf, "Discrete Choice with Social Interactions," *Review of Economic Studies*, 2001, *68* (2), 235–260.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou, "Peer Effects and Social Networks in Education," *Review of Economic Studies*, 2009, *76* (4), 1239–1267.

Card, D. and L. Giuliano, "Peer Effects and Multiple Equilibria in the Risky Behavior of Friends," *Review of Economics and Statistics*, 2013, *95* (4), 1130–1149.

Chandrasekhar, A., "Econometrics of Network Formation," in Y. Bramoullé, A. Galeotti, and B. Rogers, eds., *Oxford Handbook on the Econometrics of Networks*, 2016.

Daskalakis, C. and C. Papadimitriou, "Computing Pure Nash Equilibria in Graphical Games via Markov Random Fields," in "Proceedings of the 7th ACM Conference on Electronic Commerce" ACM 2006, pp. 91–99.

Daskalakis, Constantinos, Paul W Goldberg, and Christos H Papadimitriou, "The Complexity of Computing a Nash Equilibrium," *SIAM Journal on Computing*, 2009, *39* (1), 195–259.

de Matos, M. Godinho, P. Ferreira, and D. Krackhardt, "Peer Influence in the Diffusion of the iPhone 3G over a Large Social Network," *Management Information Systems Quarterly*, 2014, *38* (4), 1103–1133.

Fafchamps, M. and F. Gubert, "The Formation of Risk-Sharing Networks," *Journal of Economic Development*, 2007, *83*, 326–350.

González, F., "Collective Action in Networks: Evidence from the Chilean Student Movement," *PUC-Chile working paper*, 2017.

Graham, B., "An Econometric Model of Network Formation with Degree Heterogeneity," *Econometrica*, 2017, *85* (4), 1033–1063.

Granovetter, M., "Threshold Models of Collective Behavior," *American Journal of Sociology*, 1978, *83* (6), 1420–1443.

Harris, K. and R. Udry, "National Longitudinal Study of Adolescent to Adult Health (Add Health), 1994—2008 [Public Use]," *Ann Arbor, MI: Carolina Population Center, University of North Carolina-Chapel Hill [distributor], Inter-University Consortium for Political and Social Research [distributor]*, 2018.

Herings, P. and R. Peeters, "Homotopy Methods to Compute Equilibria in Game Theory," *Economic Theory*, 2010, *42* (1), 119–156.

Hoxby, C. and G. Weingarth, "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects," *Stanford working paper*, 2005.

Jackson, M., *Social and Economic Networks*, Princeton University Press, 2010.

Janson, S., T. Luczak, and A. Rucinski, *Random Graphs*, Vol. 45, John Wiley & Sons, 2011.

Jia, P., "What Happens when Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica*, 2008, *76* (6), 1263–1316.

Judd, K., P. Renner, and K. Schmedders, "Finding all Pure-Strategy Equilibria in Games with Continuous Strategies," *Quantitative Economics*, 2012, *3* (2), 289–331.

Kearns, M., "Graphical Games," in N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, eds., *Algorithmic Game Theory*, Cambridge University Press, 2007, chapter 5.

Kleinberg, J. and E. Tardos, *Algorithm Design*, Pearson Education, Inc., 2006.

Leung, M., "Inference in Models of Discrete Choice with Social Interactions Using Network Data," *arXiv preprint arXiv:1911.07106v1*, 2019.

_ **and R. Moon**, "Normal Approximation in Large Network Models," *arXiv preprint arXiv:1904.11060v1*, 2019.

Manski, C., "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 1993, *60* (3), 531–542.

Mele, A., "Does School Desegregation Promote Diverse Interactions? An Equilibrium Model of Segregation within Schools," *JHU working paper*, 2019.

Menzel, K., "Inference for Games with Many Players," *Review of Economic Studies*, 2016, *83* (1), 306–337.

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

Mode, C., *Multitype Branching Processes: Theory and Applications*, Vol. 34, American Elsevier Pub. Co., 1971.

Schelling, T., *Micromotives and Macrobehavior*, Norton & Company: New York, 1978.

Sheng, S., "A Structural Econometric Analysis of Network Formation Games," *UCLA working paper*, 2016.

Soetevent, A. and P. Kooreman, "A Discrete-Choice Model with Social Interactions: with an Application to High School Teen Behavior," *Journal of Applied Econometrics*, 2007, *22* (3), 599–624.

Turova, T., "Asymptotics for the Size of the Largest Component Scaled to "$\log n$" in Inhomogeneous Random Graphs," *Arkiv för Matematik*, 2012, *51* (2), 371–403.

Xu, H., "Social Interactions in Large Networks: A Game Theoretic Approach," *International Economic Review*, 2018, *59*, 257–284.

Xu, X. and L. Lee, "Estimation of a Binary Choice Game Model with Network Links," *OSU working paper*, 2015.