Nash Equilibria on (Un)Stable Networks*†

Anton Badev‡

6/21/2020

Abstract. In response to a change, individuals may choose to follow the responses of their friends or, alternatively, to change their friends. To model these decisions, consider a game where players choose their behaviors and friendships. In equilibrium, players internalize the need for consensus in forming friendships and choose their optimal strategies on subsets of k players - a form of bounded rationality. The k-player consensual dynamic delivers a probabilistic ranking of a game's equilibria, and, via a varying k, facilitates estimation of such games.

Applying the model to adolescents' smoking suggests that: (a.) the response of the friendship network to changes in tobacco price amplifies the intended effect of price changes on smoking, (b.) racial desegregation of high-schools decreases the overall smoking prevalence, (c.) peer effect complementarities are substantially stronger between smokers compared to between non-smokers.

Keywords: Games on Endogenous Networks, Adolescent Smoking, Multiplicity.

*The latest version of the paper, an online appendix with robustness analysis, and the implementation code are available at `www.antonbadev.net/neks`.

†Based on my Ph.D. dissertation (Badev, 2013) at the University of Pennsylvania under the guidance of Kenneth Wolpin, George Mailath and Petra Todd. I have benefited from discussions with Steven Durlauf, Hanming Fang, James Heckman, Matt Jackson, Ali Jadbabaie, Michael Kearns, Angelo Mele, Antonio Merlo, and Àureo de Paula, and from audiences at Bocconi, Cornell, 2012 SSSI (UChicago), 2012 QME (Duke), 2012 XCEDE, GWU, 2015 NSF-ITN (Harvard), Mannheim, Minnesota, 2012 NASM, NYU, Penn, Pitt/CMU, St. Louis (Econ and Olin), 2015 ESWC (Montreal), 2016 SITE (Stanford), 2014 SED (Toronto), 2014 EMES (Toulouse), Tilburg, Texas Tech, University of Hawaii, 2015 IAAE, 2016 AMES (Kyoto), and Yale (SOM) for useful comments. I gratefully acknowledge financial support from the TRIO (PARC/Boettner/NICHD) Pilot Project Competition. This work used the XSEDE, which is supported by National Science Foundation grant number OCI-1053575. All errors are mine.

‡The views expressed herein are those of the author and not necessarily those of the Board of Governors of the Federal Reserve System.

1 Introduction

In response to a change in their environment, individuals may choose to follow the responses of their friends or, alternatively, choose to change their friends. In the context of evaluating public policies (e.g., an excise tax on tobacco consumption), these decisions motivate a shift from questions such as how the friendship network *propagates* changes in individuals' behaviors, say, due to a policy change (e.g., an increase in tobacco price), to questions such as how the friendship network *responds* to such changes in individuals' behaviors. This paper studies this shift in perspective from both a theoretical and public policy view.

In order to do so, consider an environment where individuals choose both their behaviors and friendships. While these choices are fundamentally different, their difference is not related to the presence of strategic incentives or instincts for selfish decisions. Rather, choosing a friend presumes a consent (Jackson and Wolinsky, 1996) while choosing behaviors does not (Nash, 1950). The tension between the instinct for selfish choices and the consensual nature of humans' friendships can be prototyped as a game of link and node statuses where players' decision problem is augmented with a set of *stability constraints*. These constraints reflect that a player internalizes the the need for consent in forming links, or in other words, a player chooses her links only among those who desire to be her friends.

Given a player's incentives, her observed friendship links and behaviors are likely to compare favorably against her alternatives, i.e., are likely to be robust against a set of feasible deviations. Reasoning about the complexity of individuals' decision problem[1] motivates a family of equilibria indexed by the *radius* of permissible deviations. For a fixed parameter k, a Nash equilibrium in a k-stable (NEkS) network emerges when no player has profitable deviation that is permissible by the stability constraints and that involves less than k links. A feature of the proposed model is that NEkS networks are pairwise stable and, for $k = n$, NEkS networks are pairwise-Nash networks (see Jackson, 2005 for an overview of these concepts).

A primitive of games of links and behaviors is payoff externalities which can lead

[1] There are 2^{n-1} possible link deviations and only $n - 1$ possible one-link deviations.

to multiplicity of NEkS networks. To reconcile this multiplicity, this paper introduces a k-player consensual dynamic (kCD)—a family of myopic dynamic processes where players sequentially adapt their behaviors and at most $k-1$ of their links, of course, subject to the stability constraints[2]. In the presence of random preference shocks, kCDs induce a unique, invariant to the choice of k, stationary distribution over the set of all possible outcomes. Intuitively, each NEkS network is a local mode of this probability distribution. In addition, the larger k is, the faster a kCD approaches the stationary distribution.

These properties of kCDs facilitate both estimation of and simulation from these games. The model's likelihood is given by the (unique) stationary distribution of the kCD family. This distribution pertains to the Exponential Random Graph Models (Frank and Strauss, 1986; Wasserman and Pattison, 1996), for which both direct estimation and simulating from the model with known parameters are computationally infeasible.[3] The double Metropolis-Hastings sampler offers a Bayesian estimation strategy that nevertheless relies on simulations from the stationary distribution via Markov chains (Murray et al., 2006; Liang, 2010; Mele, 2017). While, for different ks, kCDs have different convergence properties they have the same stationary distribution, which in turn suggests a transparent strategy for designing these Markov chains with *varying* k.[4]

The model is estimated with data on smoking behavior, friendship networks, and home environment (parental education background and parental smoking behavior) from the National Longitudinal Study of Adolescent Health.[5] This is a longitudinal study of a nationally representative sample of adolescents in the United States, who were in grades 7–12 during the 1994–95 school year.

[2]Similar dynamics, although in a different context, are analysed by the evolutionary game theory and individual learning literatures, e.g., Foster and Young (1990); Kandori et al. (1993); Blume (1993); Jackson and Watts (2001, 2002). In a crude form, the motivation for these dynamics is present in Cournot (1838, Chapter VII) and (Nash, 1950, Section 9).

[3]A likelihood evaluation involves summations with $2^{(n^2+n)/2}$ terms, e.g. for $n = 10$, 2^{55} terms.

[4]Poor convergence properties are associated with local Markov chains, where each update is of size $o(n)$ (Bhamidi et al., 2011). Importantly, varying k on the support $\{2, \ldots, n-1\}$ is not anymore a local Markov chain. I thank an anonymous referee for pointing this out.

[5]Details about the Add Health data, including the sample construction, are in the appendix.

The empirical models in Nakajima (2007) and Mele (2017) inspired the proposed framework. Nakajima (2007) studies peer effects abstracting from friendships and Mele (2017) obtains large network asymptotics of a model with link formation only. The approaches in these papers are fundamentally compatible so these models can be unified in a joint model, as in Boucher (2016) and Hsieh et al. (2016). Compared to existing empirical frameworks (including Canen et al., 2016; Boucher et al., 2019; Battaglini et al., 2019), this paper explicitly models the strategic incentives guaranteeing network stability in the sense of Jackson and Wolinsky (1996).

The empirical analysis of friendship networks and smoking behaviors lends support to a host of results which are related to the large body of empirical work on social interactions and teen risky behaviors. Typically, empirical studies on peer effects either lack data on friendship network or take the friendship network as given.[6] Also the approaches range from models that directly relate an individual's choices to mean characteristics of his peer groups (e.g., see Powell et al. (2005) and Ali and Dwyer (2009a)) to models with elaborate equilibrium micro-foundations, such as those in Brock and Durlauf (2001, 2007); Krauth (2005); Calvó-Armengol, Patacchini and Zenou (2009). In terms of estimates, this paper makes the first step in explaining how not accounting for the response of these social network could bias the estimates.[7] Similarly, this paper pioneers a mechanism capable of explaining the role of the school composition, or more generally the determinants of the social fabric, on teen risky behaviors. The possibility of such a role was theorized by Graham et al. (2014) and experimentally discovered by Carrell et al. (2013).

1.1 Conclusions from the empirical analysis

The model is estimated under various restrictions on the parameter specification and on the data availability. Two observations merit noting at the estimation stage.

[6]See, for example, Liu et al. (2014), who distinguish between local aggregate and local average peer effects, and the references therein.

[7]It is difficult, if not impossible, to account the empirical contributions of the large literature on peer effects and teen risky behaviors. For a small sample of papers obtaining estimates of peer effects see Chaloupka and Wechsler (1997), Ali and Dwyer (2009b) and the references in (CDC, 2000, Surgeon General's Report).

First, the peer effect complementarities are substantially stronger between smokers compared to between non-smokers. The model's parametrization permits differentiating these externalities and the conclusion is notable from the parameter estimates. Second, lack of network data, which forces the estimation to suppress the local peer effect externalities, substantially biases downwards the price coefficient.

The obtained sets of estimates are used to perform counterfactual experiments under various estimation scenarios. The purpose of these experiments is to quantify the response of the friendship network to policies targeting adolescent smoking. A by-product of this analysis is an assessment of the bias in the model's predictions due to lack of network data or due to various miss-specifications.

The first experiment asks whether this response is relevant for policies working through changes in tobacco prices. To motivate this exercise, compare how individuals respond to a price increase in fixed versus endogenous network environments. There are two effects to consider. The *direct* effect of changing tobacco prices is the first order response and, intuitively, will be larger whenever individuals are free to change their friendships. That is, more individuals are likely to immediately respond to changes in tobacco prices provided they are not confined to their (smoking) friends. The *indirect* (ripple) effect of changing tobacco prices is the effect on smoking which is due, in part, to the fact that one's friends have stopped smoking. Contrary to before, a fixed network propels further the indirect effect. In a fixed network, an individual who changes her smoking status is bound to exert pressure to her (fixed) friends who are most likely smokers. It is, then, an empirical question how these two opposing effects balance out. Simulations with the full model and with a model where the friendship network is kept fixed suggest that the direct effect dominates. In other words, following an increase in tobacco prices the response of the friendship network amplifies the intended reduction in smoking prevalence.

The second experiment asks whether school racial composition has an effect on adolescent smoking. When students from different racial backgrounds study in the same school, they interact and are likely to become friends. Being from different racial backgrounds students have different intrinsic propensity to smoke and the question is what is the equilibrium behavior in these mixed-race friendships: do those who do not smoke start smoking or those who smoke stop smoking? Simulations from the model suggest that redistributing students from racially segregated schools into racially balanced schools decreases the overall smoking prevalence.

The last experiment simulates a small scale policy intervention targeting only a part of school's population. The policy is efficient so that those exposed to the treatment stop smoking. At the same time it is not feasible (too costly) to treat the entire school. In this experiment, the question is when treated individuals return, will their friends follow their example, i.e. extending the effect of the proposed policy beyond the set of treated individuals and thus creating a domino effect, or will their pre-treatment friends un-friend them? In essence, this is a question about the magnitude of the spillover effects and this study suggests that aggregate spillovers are roughly double compared to the scale of the policy.

1.2 Related literature

This paper studies and estimates a game on endogenous network where players choose both their behaviors (e.g., smoking) and friendship links. The proposed model can be restricted to a game played on a fixed network. These games date back to the physics literature of the 70s and in economics have been analyzed with both discrete and continuous choices (e.g., see Jackson and Zenou, 2015 and Bramoullé and Kranton, 2016 for surveys). Most of the empirically tractable games have been developed either in continuous settings (e.g., Ballester et al., 2006, Bramoullé et al., 2014, Calvó-Armengol et al., 2009) or, when data on the friendship network is not available, restricting the model further to where peer effects are measured via group averages (e.g., see Brock and Durlauf, 2001, 2007, Nakajima, 2007, and the survey in Blume et al., 2015).

Symmetrically, the proposed model can be restricted to a network formation game (e.g., see Jackson (2008) for a systematic textbook presentation). A large and growing body of studies on the economics of these games followed Jackson and Wolinsky (1996) who, in a departure from the traditional non-cooperative game paradigm, introduced the notion of pairwise stability. In this paper, the stability constraints guarantee that any NEkS play is pairwise stable and for $k = n$ such play is pairwise-Nash (see Myerson, 1991; Calvó-Armengol, 2004; Goyal and Joshi, 2006; Bloch and Jackson, 2006, 2007 and the survey in Jackson, 2005).

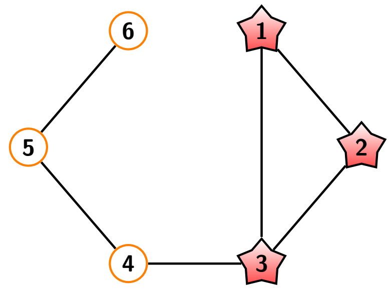
A handful of theoretical papers consider both network formation along with other choices potentially affected by the network (see Goyal and Vega-Redondo, 2005; Cabrales et al., 2011; König et al., 2014; Baetz, 2015; Lagerås and Seim, 2016; Hiller, 2017; Jackson, 2018). Importantly, the theoretical frameworks available are meant to provide focused insights into isolated features of networks and deliver sharp predictions, while abstracting from players' heterogeneity and so are not easily adapted for the purposes of estimation.

Econometric models of networks and actions are proposed in Goldsmith-Pinkham and Imbens (2013), Hsieh and Lee (2016) and Johnsson and Moon (2017) where the decisions to form friendships influence the decision to engage in a particular activity. The focus of their research, however, is not on policy analysis nor on accounting for the possible endogenous response of the friendship network to changing the decision environment. In contrast, the framework proposed in Boucher (2016) is microfounded as a particular equilibrium in a non-cooperative model of friendships and behaviors. Related work by Hsieh et al. (2016) proposes a two-stage estimation procedure, with an application to R and D, which relies on conditional independence of links delivered by abstracting from link externalities. Canen et al. (2016) propose an empirically tractable framework, building on Cabrales et al. (2011), where politicians choose both socialization and legislation efforts, and study bill cosponsorship in the U.S. Congress.[8] Differently to this literature, the proposed model is founded on the explicite strategic incentives that guarantee network stability in the sense of Jackson and Wolinsky (1996).

Finally, adaptive dynamic and potential function representation, as a dimension-

[8]There are recent contributions to the econometrics literature which focus on link formation, though these are not easily extendable to include action choice as well, e.g., see Sheng (2016), Chandrasekhar and Jackson (2016), Leung (2014), de Paula et al. (2018), Graham (2017), Menzel (2015) and the reviews in Chandrasekhar (2015); de Paula (2016); Bramoullé et al. (2016).

Figure 1: An example



Note: In the graph, each player in $I = \{1, 2, 3, 4, 5, 6\}$ is depicted as a vertex and a friendship is depicted as an edge, e.g. 1 and 2 are friends. The star-shaped shaded nodes denote players who smoke tobacco, e.g. 1, 2, and 3 are smokers.

ality reduction tool, are widely used in (algorithmic) game theory, computer science and in economics of networks for processes on fixed networks, for processes of link formation and, more recently, for combined processes, e.g. Foster and Young (1990), Blume (1993), Jackson and Watts (2001, 2002), Nakajima (2007), Bramoullé et al. (2014), Bourlés et al. (2017), Mele (2017), Boucher (2016), Hsieh and Lee (2016). In contrast to this literature, this paper highlights a slightly different role of the potential function, namely, as a tool to justify the gravitation of a family of adaptive dynamics around the equilibria of the static links and behaviors game. Further, the analysis of the kCD family justifies a model-based approach to simulate from and estimate these processes.

2 A game on an endogenous network

Imagine a world populated by individuals who chose both their friends and their behaviors, e.g. to smoke or not. Figure 1 provides an example with 6 individuals. In the figure, individuals are depicted as nodes on a graph and the star-shaped shaded nodes are those who smoke. Next, friendships are depicted as links between pairs of nodes. These links are undirected because (being in) a friendship is a symmetric binary relation.

Before introducing the details of the game, it is worth enumerating the distinct features of this decision environment that are also reflected in the model. First, player i's choices of behavior and friendships are different in that friendships (unlike behaviors) require consent to form and maintain. Second, there are likely to be externalities not only between individual's behavior and the behaviors her friends but also between individuals' friendship decisions. These various types of externalities are explicitly defined in the proposed payoff specification. Finally, this is a complex decision environment in that even with 10 individuals, each player considers roughly 1000 alternative strategies while with 100 individuals each player considers roughly 10^{30} alternative strategies. This complexity relates to the proposed (family of) equilibria and adaptive dynamic.

The model is developed in two stages. First, agents' strategic behavior is analyzed in static settings. Then, section 3 develops a family of myopic dynamic processes used to approximate the predictions of the static model in a inferentially convenient way.

2.1 Players and preferences

Each i, in a finite population $I = \{1, 2, ..., n\}$, chooses $a_i \in \{0, 1\}$ and a set of links $g_{ij} = g_{ji} \in \{0, 1\}$ for $j \neq i$. In the settings of adolescents' smoking and friendship decisions, I is the set of all students in a given high school, $a_i = 1$ if student i smokes, and $g_{ij} = 1$ if i and j are friends. These are the settings in figure 1 above, with 3 smokers, e.g. $a_1 = a_2 = a_3 = 1$, and 6 friendships, e.g. $g_{12} = 1$, $g_{23} = 1$, etc. A final piece of the description of the population is individuals' exogenous characteristics X_i, e.g. age, gender, etc.

Individual i chooses her behavior and friendship statuses $S_{(i)} = (a_i, \{g_{ij}\}_{j \neq i})$ from her choice set $\mathbf{S}_{(i)} = \{0, 1\}^n$ to maximize her payoff u_i. Let $S = (S_{(1)}, \ldots, S_{(n)}) \in \prod_i \mathbf{S}_{(i)} = \mathbf{S}$ and $X = (X_1, ..., X_n) \in \mathbf{X}$. Formally i's payoff function, $u_i : \mathbf{S} \times \mathbf{X} \longrightarrow \mathbb{R}$, orders the outcomes in $\mathbf{S}$ given X:

$$u_i(S,X) = a_i v_i + \underbrace{a_i \phi \sum_{j \neq i} a_j}_{\text{aggr. externalities}} \tag{1}$$

$$+ \underbrace{\phi_S \sum_j g_{ij} a_i a_j + \phi_N \sum_j g_{ij}(1-a_i)(1-a_j)}_{\text{local externalities}} \tag{2}$$

$$+ \sum_j g_{ij} w_{ij} + \underbrace{q_{ijk} \sum_{\substack{j,k \\ j<k}} g_{ij} g_{jk} g_{ki}}_{\text{clustering}} - \psi \underbrace{\left(\frac{1}{2}(d_i^2 + d_i) + \sum_{j \neq i} g_{ij} d_j \right)}_{\text{(convex) } cost_i(d_i, \{d_j\}_{j \in d_i})} \tag{3}$$

where $d_i = \sum_j g_{ij}$ is the degree (total number of links) of i. Here $v_i = v(X_i)$, $w_{ij} = w(X_i, X_j)$ and $q_{ijk} = q(X_i, X_j, X_k)$ are functions of agents' (exogenous) characteristics. To avoid clutter in the summation ranges, assume that g_{ii} is defined and equal to zero for all i so that, for example, $d_i = \sum_{j \neq i} g_{ij} = \sum_j g_{ij}$.

Note how u_i depends both on individual's exogenous characteristics X_i (e.g., terms v_i and w_{ij}) and on her endogenous characteristics, e.g. number of friends, smoking statuses of her friends, and etc. More specifically, the terms in payoff (1-3) can be sorted into three groups: terms that relate to the incremental payoff of changing a_i, terms that relate to the incremental payoff of changing g_{ij} and terms that relate to both.

The first three terms in (1-3) relate to the incremental payoff of changing i's behavior a_i conditional on the friendship network,

$$\Delta_{a_i} u_i(S,X) = v_i + \phi \sum_{j \neq i} a_j + \phi_S \sum_{j \neq i} g_{ij} a_j - \phi_N \sum_{j \neq i} g_{ij}(1-a_j).$$

The first term v_i is the (exogenous) intrinsic utility of choice $a_i = 1$ which is allowed to vary with i's attributes X_i. The second term $\phi \sum_{j \neq i} a_j$ captures the aggregate externalities. That is, i may be influenced from the behaviors of the surrounding population $\sum_{j \neq i} a_j$, provided $\phi \neq 0$. The last two terms in $\Delta_{a_i} u_i(S,X)$ are the differential of the local externalities $\phi_S \sum_j g_{ij} a_i a_j + \phi_N \sum_j g_{ij}(1-a_i)(1-a_j)$ in (2).

Note that $a_i a_j$ equals 1 if and only if $a_i = a_j = 1$ so that, conditional on the friendship network, this term captures pressures on i to follow (or to break away if $\phi_S < 0$) her friends' decision to chose 1 (to smoke). Analogously, $(1 - a_i)(1 - a_j)$ equals 1 if and only if $a_i = a_j = 0$, and this term captures pressures on i to conform to the behaviors of her choosing 0 (non-smoking) friends. Because ϕ_S need not equal ϕ_N, the opposing conformity pressures from friends who choose 1 and from friends who choose 0 need not be equal in magnitude. Finally, as will become evident shortly, the local externalities terms are related to the incremental payoff of changing g_{ij} where, conditional on individuals' actions, these terms capture a tendency to befriend others playing the same action. To sum up, an agent's utility increases by ϕ_S with every friend who plays the same action if that action is 1, and by ϕ_N with every friend who plays the same action if that action is 0.

The last four terms in (1-3) relate to the incremental payoff to i of changing g_{ij} conditional on players' actions:

$$\Delta_{g_{ij}} u_i(S, X) \quad = \quad w_{ij} + q_{ijk} \sum_k g_{ik} g_{jk} - \psi(d_i + d_j) + \phi_S a_i a_j + \phi_N (1 - a_i)(1 - a_j).$$

The first term w_{ij} captures the (exogenous) utility of a friendship which may depend on i's and j's degree of similarity, i.e., same sex, gender, race, etc. The next term is the differential of $q_{ijk} \sum_{j<k} g_{ik} g_{jk} g_{ki}$ in (3) which captures link externalities. Mechanically, i may have preferences for whether or not her friends are friends themselves. In particular, i may prefer sharing her friends ($q > 0$) or, on the contrary, prefer friendship exclusivity ($q < 0$).[9] The third term is the differential of the convex cost term in (3) which reflects the costs of establishing a friendship between i and j. Properties of the cost term to note are: (i.) the more friends i has, the more costly it is for i to establish an additional friendship and (ii.) the costs are shared so for i it is more costly to maintain friendships with more popular (high d_j) as opposed to less

[9] A compelling interpretation of this term is consistent with the presence of meeting frictions. In particular, meeting and befriending friends of friends can explain the tendency of individuals to form triangles of friendships (e.g., see Jackson and Rogers, 2007). This paper studies relatively small friendship networks so frictions are less likely to play a pronounced role.

popular (low d_j) individuals. The last two terms relate to the previously discussed local externalities terms $\phi_S \sum_j g_{ij} a_i a_j + \phi_S \sum_j g_{ij}(1-a_i)(1-a_j)$ in (2).

2.2 Equilibrium play

Given a player's preferences, her observed links and action are likely to compare favorably against her alternatives. However, the number of available alternatives renders players' decision problem complex[10] and motivates a family of equilibria where players consider only strategies that are close by, or in other words, where players consider deviations that are small. Here, the notion of closeness naturally translates to strategies that involve changing only few links. A final point concerning the equilibrium play is that players are aware that links are formed with consent.

Definition 1 *A profile of actions and a network* $S^* = (\{a_i^*\}_{i\in I}, \{g_{ij}^*\}_{i\in I, j\in I\backslash i})$ *is a* **Nash equilibrium in a k(-player) stable (NEkS) network**, *provided* $S^*_{(i)} = (a_i^*, \{g_{ij}^*\}_{j\neq i})$ *is a solution of* i'*s decision problem on* $I_k \subseteq I$*:*

$$\max_{a_i, \{g_{ij}\}_{j\in I_k\backslash i}} \quad u_i(a_i, \{g_{ij}\}_{j\in I\backslash i}; S^*_{-i}) \tag{4}$$

$$\text{s.t.} \quad g_{ij} = 1 \quad \text{only if} \quad \Delta_{g_{ij}} u_j(a_i, \{g_{ij}\}_{j\in I\backslash i}; S^*_{-i}) \geq 0 \quad \forall j \in I_k\backslash i \tag{5}$$

where $1 < k \leq n$*,* $I_k = \{i\} \cup \{i_1, \ldots, i_{k-1}\}$ *and* $i \notin \{i_1, \ldots, i_{k-1}\}$*, for all* i *and* I_k*.*

To state the above definition in words, in a NEkS network no player has permissible, by the stability constraints (5), and profitable deviation involving changing the statuses of less than k links. A notable feature of the NEkS networks is that not only links are formed with consent but also players internalize the need for consent through subjecting their play to the stability constraints. The stability constraints owe their name to their relation to the notion of stability introduced in Jackson and Wolinsky (1996) (see proposition 2 below).

Assumption 1 *Assume that* $w()$ *and* q *are symmetric in their arguments/indices.*

[10] There are 2^{n-1} possible link deviations and only $n-1$ possible one-link-at-a-time link deviations.

Proposition 1 *With the utilities in* (1)

1. *For any S, k, i and I_k, the problem in (4-5) is well defined and has a solution.*

2. *For any k, a NEkS network exists.*

The existence of a solution to the individual's decision problem in (4-5) and an equilibrium follows from the existence of potential function for this game (Monderer and Shapley, 1996). The proof is in appendix A (p. 33).

Proposition 2 *With the utilities in* (1)

1. *For $k = 2$, NEkS networks are pairwise stable;*

2. *For $k = n$, NEkS networks are pairwise-Nash networks;*

3. *For $k' < k$, any NEkS network is also a NEk'S network.*

Part 1 can be strengthen for any preferences: for $k = 2$, any NEkS play is pairwise stable (Jackson and Wolinsky, 1996). For $k = n$, NEkS networks are pairwise-Nash networks (Calvó-Armengol, 2004; Goyal and Joshi, 2006; Bloch and Jackson, 2006, 2007). Finally, the NEkS family is ordered by set inclusion so that the existence of a pairwise stable network is a necessary condition for the existence of a NEkS network. The proof is in appendix A (p. 34).

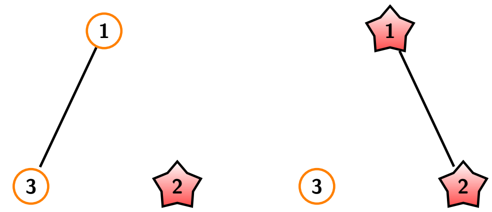
2.3 An example

To see how the choice of k may affect the equlibrium networks, consider a simplified version of payoffs (1-3) where all externalities other than the local peer effects and costs are absent. Let $I = \{1, 2, 3\}$, $\phi = 0$, $\phi_S = \phi_N = \phi_0$, and $q = 0$. Also let $w_{ij} = 0$ for all i and j so that

$$u_i = v_i + \phi_0 \sum_j g_{ij}(a_i a_j + (1 - a_i)(1 - a_j)) - \psi \underset{i}{cost}(d_i, \{d_j\}_{j \in d_i}). \tag{6}$$

Further, let $v_2 = \bar{v}$ and $v_3 = -\bar{v}$ for $\bar{v}$ large so that it is always a dominant strategy for players 2 and 3 to choose $a_2 = 1$ (smoke) and $a_3 = 0$ (not smoke) respectively.

Figure 2: Examples of NEkS networks for $k = 2$ and $k = 3$



Note: Let $I = \{1, 2, 3\}$ and $u_i = v_i + \phi \sum_j g_{ij}(a_i a_j + (1 - a_i)(1 - a_j)) - \psi\, cost_i(d_i, \{d_j\}_{j \in d_i})$. It is straightforward to construct examples where for $k = 3$ there is a unique NEkS network while for $k = 2$ both of the depicted networks are NEkS networks.

Finally, if it is costly to acquire friends ($\psi > 0$) then players will never choose a friend playing different action because $w_{ij} = 0$, and if the benefits of having a friend that plays the same action outweigh these costs ($\phi_0 > 2\psi$) then player 1 would want to have exactly one friend (either player 2 or player 3) with the same smoking status. These candidates for equilibrium are depicted in figure 2.

For $k = 3$ there is (generically) a unique NEkS network. If $v_1 < 0$ then player 1 chooses not smoke and befriends player 2 (figure 2 left) else ($v_1 > 0$) player 1 chooses to smoke and befriends player 3 (figure 2 right). In contrast, for $k = 2$ if $v_1 \in (-\phi_0 + 2\psi, \phi_0 - 2\psi)$ both networks in figure 2 are NEkS networks. Note that the larger the complementarities are (ϕ_0), the larger the region for v_1 is where there are multiple NEkS networks.

3 Consensual dynamic. An estimable framework

The NEkS play offers an intuitive prescription for the *outcomes* of the forces driving behaviors and friendships, without specifying the decision *process* leading to these outcomes. This abstraction is challenged by strong informational assumptions where players are presumed to correctly anticipate other players' choices and by the presence of multiple NEkS networks none of which can be ruled out a priori. Turning to a framework based on adaptive dynamics and random utility delivers a way to embed this multiplicity into an inferentially convenient framework.

Formulation (4-5) of individuals' decision problem provides a basis for an adaptive process where behaviors and friendships evolve towards (or around) a NEkS network. The general idea that equilibrium might arise from simple (myopic) adaptive dynamic as opposed to from a complex reasoning process is very intuitive. The particular emphasis, compared to the interpretations in Kandori et al. (1993), Blume (1993) and Jackson and Watts (2001, 2002), is on obtaining an estimable structure via a flexible adaptive process (parametrized via k).[11] This flexibility presents advantages in simulating and estimating these games.

3.1 k-player consensual dynamic (kCD)

Every period $t = 1, 2, \ldots$ a randomly chosen individual, say i, considers $k-1$ of her friendships, say with $\{i_1, \ldots, i_{k-1}\}$, and her behavior a_i. In particular, i myopically solves her decision problem (4–5) on $I_k = \{i\} \cup \{i_1, \ldots, i_{k-1}\}$. A stochastic meeting process μ_t outputs i and I_k:

$$\Pr\left(\mu_t = (i, I_k)\,|S_{t-1}, X\right) = \mu_{i,I_k}\left(S_{t-1}, X\right) \tag{7}$$

In the simplest case, when any meeting is equally probable, $\mu_{i,I_k}\left(S_{t-1}, X\right) = \frac{1}{n}\frac{1}{\binom{n-1}{k-1}}$ for all i, I_k, S_{t-1}, and X. However, we only need that any meeting is possible.

Assumption 2 $\mu_{i,I_k}\left(S_{t-1}, X\right) > 0$ *for all* $i \in I$, I_k, $S \in \mathbf{S}$ *and* $X \in \mathbf{X}$.

The sequence of meetings together with players' optimal decisions induce a sequence of network states (S_t), which is indexed by time subscript t and which will be referred to as *k(-player) consensual dynamic* (kCD).

Proposition 3 *Fix $k \in [2, n]$. With assumptions 1 and 2, for a kCD S_t:*

1. *Any NEkS network is absorbing, i.e. $S_{t'} = S_t$ if S_t is a NEkS network, $t' > t$;*

[11] The literature on stochastic stability has studied stochastic dynamic where shocks vanish over time as a tool for equilibrium selection. In addition, a typical approach has been to analyze adaptive dynamics where either agents take turns to update their strategies (i.e., in our settings, all links) or a pair of players update the status of their link.

2. Independently of the initial state $\Pr(\lim_{t\to\infty} S_t \in NEkSN) = 1$.

Indeed, for any k, the NEkS networks are exactly the rest points of simple decision processes, the kCDs. The proof is in Appendix A (p. 34).

3.2 kCDs with random utility

Consider the following modification of a kCD. In each period after the meeting is realized, the decision problem of the player who is drawn to make a choice is cast as a random utility choice.[12] That is, player's payoffs for each alternative are augmented with a random component, ultimately making her choice stochastic. Because players' choices are stochastic, such a kCD with random utility delivers a distribution over possible NEkS networks as opposed to a single network (see proposition 3). Moreover, this distribution has some convenient properties when treated as (the) likelihood.

Assumption 3 *Suppose that the utilities in (1) contain an additive random preference shock* $u_i(S, X) + \epsilon_S$ *where* $\epsilon_S \sim$ *i.i.d. across time and network states. Moreover, suppose that* ϵ_S *has c.d.f. and unbounded support on* $\mathbb{R}$.

Assumption 4 *Suppose that the preference shock* ϵ *is distributed Gumbel*$(\mu_\epsilon, \beta_\epsilon)$.

Assumption 5 *Suppose that the meeting probability in* (7), $\mu_{i,I_k}(S, X)$ *does not depend on* a_i *and* g_{ij} *for all* $j \in I_k$. *(Alternatively, which is slightly weaker, suppose that* $\mu_{i,I_k}(S, X) = \mu_{i,I_k}(S', X)$ *for all* $S, S' \in \mathbf{S}$.*)*

The meeting process $\{\mu_t\}_{t=1}^{\infty}$ and the sequence of optimal choices, in terms of behaviors and friendship links, induce a Markov chain on $\mathbf{S}$ referred to as a kCD with random utility. The family of kCDs with random utility obey some desirable properties. (The proof is on p. 35.)

Theorem 1 [Stationary distribution] *Fix* $k \in [2, n]$. *The* k*CD with random utility has the following properties:*

[12] This is also known as the random utility model. See Thurstone (1927), Marschak (1960), McFadden (1974) and, for textbook treatement, (Train, 2003, Chapter 2).

1. *With assumptions 2 and 3, there is a unique stationary distribution* $\pi_k \in \Delta(\mathbf{S})$ *for which* $\lim_{t\to\infty} \Pr(S_t = S) = \pi_k(S)$. *In addition, for any function* $f : \mathbf{S} \to \mathbb{R}$, $\frac{1}{T}\sum_{t=0}^{T} f(S_t) \longrightarrow \int f(S)\, d\pi_k$.

2. *With assumptions 1-5,*

$$\pi(S,X) \propto \exp\left(\frac{\mathcal{P}(S,X)}{\underline{\beta}}\right). \tag{8}$$

In particular, $\pi(S,X)$ *does not depend on* k.

The first part is not surprising in that it asserts that a kCD with random utility is well behaved so that standard convergence results apply. The uniqueness of π_k precludes dependence between snapshots from this process and its initial state, and the ergodicity allows to simulate from π_k via drawing a long trajectory of the kCD.

The second part of the theorem has implications for implementing the model. Note how in (8) the stationary distribution π does not depend on k and, thus, delivers a tool to unify the equilibria in the NEkS family. In particular, π ranks in a probabilistic sense the family of equilibria within and across different ks (see theorem 3). This is particularly relevant for implementing the model when π can be treated as the likelihood. In addition, the expression in (8) provides for a transparent identification of model's parameters. It is clear that, given the variation in the data of individual choices $\{a_i\}_{i=1}^n$, friendships $\{g_{ij}\}_{i,j=1}^n$ and attributes $\{X_i\}_{i=1}^n$, functional forms for $v, w, q, \psi, \phi, \phi_S, \phi_N$ will be identified as long as the different parameters induce different likelihoods of the data. Finally, a closed-form expression for π facilitates the use of likelihood-based methods for estimating model's parameters.

3.3 Speed of convergence

The kCDs with random utility depend on k in an important way despite the fact that their stationary distribution is invariant to k. The next result studies this dependence in isolation from all other determinants of the kCDs with random utility.

Theorem 2 [kCDs ranking] *Set* $v_i = w_{ij} = h = \phi = \phi_S = \phi_N = q = \psi = 0$. *Then, the second eigen value of the* $2^{(n^2+n)/2}$*-by-*$2^{(n^2+n)/2}$ *transition matrix of the* k*CD is given by:*

$$\lambda_{k,[2]} = \frac{1}{n}\left(n - 1 + \frac{n-k}{n-1}\right) \tag{9}$$

In particular, $\lambda_{k',[2]} < \lambda_{k,[2]}$ *for* $2 \leq k < k' \leq n$ *so that the* k'*CD converges strictly faster than* k*CD to the stationary distribution* π.

In the hypothesis of theorem 2, all payoff parameters in equations (1-3) are set to zero so that players do not differentiate between different networks (i.e. $u_i(S; X) = 0$ for all $S \in \mathbf{S}$) which implies that the kCDs traverse in unbiased way the space of all possible networks $\mathbf{S}$. In the end, the stationary distribution π is one where the behaviors and network links are i.i.d. Poisson(0.5) and, importantly, k is the only determinant of kCDs' transition probabilities and convergence rates.[13]

There are two rationales behind pursuing a characterization of the speed of convergence of kCDs. As anticipated (and formally established shortly) π probabilistically ranks the family of NEkS networks. In a dual fashion, the differential speed of convergence provides a means to rank the family of kCDs with random utilities. In particular, the larger k is, the smaller is the second eigen value $\lambda_{k,[2]}$, i.e. the faster kCDs converge to π (see Debreu and Herstein, 1953, Section 4). In this sense, a snapshot of the state (drawn from π) is more likely to reflect a kCD with random utility where k is large as opposed to one where k is small.

The second reason for why properties of kCDs are of their own interest is highlighted by Bhamidi et al. (2011) who show that adaptive dynamic with local updates (i.e. $o(n)$ links at a time) converges very slowly. Such slow convergence rates could question the conceptual treatment of the limiting distribution π as a likelihood. For this same reason, simulation based methods that rely on local updates may not work in practice for estimation/simulation of these models.[14] Note that kCDs encompass

[13] In general, the shape of the potential, i.e. the terms of the potential function, and the geography of the network will likely influence the speed of convergence. To the best of my knowledge, treatment of the general case remains out of reach.

[14] See the discussion in Chandrasekhar and Jackson (2016).

not only local updates, e.g. $k = [n/2]$, and thus suggest a way to avoid the problem of slow convergence (poor approximation). Relatedly, theorem 2 offers insights into an important trade-off for sampling design: the Markov chain is facing a trade-off between speed of convergence and complexity in simulating the next step. For small k, the convergence to π is slower, however, the update is drawn from a discrete distribution with small (2^k) support. The opposite holds when k is large.[15]

3.4 Discussion

3.4.1 Probabilistic ranking. The most probable equilibria

The stationary distribution obtained in theorem 1 gives an intuitive (probabilistic) ranking of the family of NEkS networks. Under π, a network state will receive a positive probability, although it may not be an equilibrium in any sense. It will be desirable, however, that in the vicinity of an equilibrium, the equilibrium to receive the highest probability. Relatedly, the mode of π (i.e. the state with the highest probability) has special role. This offers a new perspective to the theoretical results on equilibrium selection from evolutionary game theory, namely equilibrium ranking.

To formalize our discussion, define the neighborhood $\mathbf{N} \subset \mathbf{S}$ of $S \in \mathbf{S}$ as:

$$\mathbf{N}(S) = \{(g_{ij}, S_{-ij}) : i \neq j, g_{ij} \in \{0,1\}\} \bigcup \{(a_i, S_{-i} : a_i \in \{0,1\}\}$$

Theorem 3 *Suppose assumptions 1-5 hold.*

1. *A state $S \in \mathbf{S}$ is a Nash equilibrium in a pair-wise stable network iff if it receives the highest probability in its neighborhood $\mathbf{N}$.*

2. *The most likely network states $S^{mode} \in \mathbf{S}$ (the ones where the network spends most of its time) are pairwise Nash networks.*

[15] The structure of the problem permits a substantial computational shortcut within the MH algorithm for generating the update of kCD. In particular, for any k computing the acceptance probability scales only quadraticly with the size of the network because it is enough to compute the change in potential as opposed to the potential itself. The published code of the paper contains more details.

3.4.2 A kCD with random k

Consider what appears to be a very unrestrictive meeting process, where every period a random individual meets a set of potential friends of random size and composition. Let κ be a discrete process with support $2, \ldots, n$ and augment the meeting process with an additional initialization step with respect to the dimension of μ. In particular, at each period first κ is realized and then μ^k is drawn just as before. It is relatively straightforward to establish, without any assumptions on the process κ, that this augmented process has the same stationary distribution π as the one from theorem 1.[16] This is another demonstration of the fact that different meeting processes result in observationally equivalent models.

4 Data and estimation

4.1 The Add Health data

The National Longitudinal Study of Adolescent Health is a longitudinal study of a sample of adolescents in grades 7–12 in the United States in the 1994–95 school year. The sample is representative of US schools with respect to region of country, urbanicity, school size, school type, and ethnicity. In total, 80 high schools were selected together with their “feeder“ schools. The students were first surveyed in-school and then at home in four follow-up waves conducted in 1994–95, 1996, 2001–02, and 2007–08. This paper makes use of Wave I of the in-home interviews with students enrolled in the schools from the so called saturated sample. Only for schools from the saturated sample, all of their students were eligible for in-home interviews.

The in-home interviews contain rich data on students’ behaviors, home environment, and friendship networks. These data are merged with administrative data on the average price of a carton of cigarettes from the American Chamber of Commerce Research Association (ACCRA). ACCRA’s data are linked to the Add Health data on the basis of state and county FIPS codes for the year in which the data were col-

[16]A formal statement and a proof are omitted because these follow the ones of theorem 1.

lected. Additional details about the estimation sample including sample construction and sample statistics are presented in the appendix.

4.2 Bayesian estimation

The kCDs with random utility deliver a unique stationary distribution π which for estimation purposes can be thought of as likelihood. Because no information is available on when the process started or on its initial state, the best prediction about the current state is given by π. For a single observation $S \in \mathbf{S}$, the likelihood is given by:

$$p(S|\theta) = \frac{\exp\{\mathcal{P}_\theta(S)\}}{H_\theta} \tag{10}$$

where $\mathcal{P}_\theta$ is the potential (evaluated at θ) and $H_\theta = \sum_{S\in\mathbf{S}} \exp\{S\}$ is an (intractable) normalizing constant.[17] The specific form of the likelihood pertains to the exponential family, whose application to graphical models has been termed as Exponential Random Graph Models (ERGM).[18]

The estimation draws from the Bayesian literature on approximating likelihoods with intractable normalizing constant developed in Murray et al. (2006) and Liang (2010). The proposed implementation augments their algorithm with an extra step informed by properties of the kCDs in the proposed model.

The posterior sampling algorithm is exhibited in table 1. In the original double M-H algorithm, an M-H sampling of S from $\pi_\theta(S)$ is nested in an M-H sampling of θ from the posterior $p(\theta|S)$. The new piece in table 1 is the random meeting process in step 5. Theorem 2 suggests that varying k improves the convergence and theorem 1 demonstrates that changing k leaves the stationary distribution unaltered. Proposition 4 below demonstrates the validity of the algorithm.

Proposition 4 [Varying double M-H algorithm] *Let $1 < k \leq n$ and suppose assumptions 2 and 3 hold. If in the algorithm of table 1, the proposal density*

[17] The summation in calculating H_θ cannot be computed directly for practical purposes even for small n, e.g., for $n = 10$ this summation includes 2^{55} terms.

[18] ERGMs are a broad class of statistical models, capable of incorporating arbitrary dependencies among the links of a network. See Frank and Strauss (1986) and Wasserman and Pattison (1996).

Table 1: Varying double M-H algorithm

Input: initial $\theta^{(0)}$, number of iterations T, size of the Monte Carlo R, data S	
1.	**for** $t = 1 \dots T$
2.	Propose $\theta' \sim q(\theta'; \theta^{(t-1)}, S)$
3.	Initialize $S^{(0)} = S$
4.	**for** $r = 1 \dots R$
5.	Draw $k \sim p_k(k)$
6.	Draw a meeting $\mu(i, I_k)$ where $i \in \{1 \dots N\}$ and $I_k \subset \{1 \dots N\} \backslash \{i\}$ from $q_\mu(i, I_k)$
8.	Propose S' where $(a_i, \{g_{ij}\}_{j \in I_k})$ are drawn from $q_\mu(S' \vert S^{(r-1)}; (i, I_k))$
9.	Compute $\bar{a} = \frac{\exp\{\mathcal{P}_{\theta'}(S')\}}{\exp\{\mathcal{P}_{\theta'}(S^{(r-1)})\}} \frac{Q(S^{(r-1)} \vert S'; p_k, q_{i,I_k})}{Q(S' \vert S^{(r-1)}; p_k, q_{i,I_k})}$
10.	Draw $a \sim$ Uniform$[0, 1]$
11.	If $a < \bar{a}$ then $S^{(r)} = S'$ else $S^{(r)} = S^{(r-1)}$
12.	**end** for $[r]$
13.	Compute $\bar{a} = \frac{q(\theta^{(t-1)}; \theta')}{q(\theta'; \theta^{(t-1)})} \frac{p(\theta')}{p(\theta^{(t-1)})} \frac{\exp\{\mathcal{P}_{\theta^{(t-1)}}(S^{(R)})\}}{\exp\{\mathcal{P}_{\theta^{(t-1)}}(S)\}} \frac{\exp\{\mathcal{P}_{\theta'}(S)\}}{\exp\{\mathcal{P}_{\theta'}(S^{(R)})\}}$
14.	Draw $a \sim$ Uniform$[0, 1]$
15.	If $a < \bar{a}$ then $\theta^{(t)} = \theta'$ else $\theta^{(t)} = \theta^{(t-1)}$
16.	**end** for $[t]$

conditional on meeting (i, I_k), $q_\mu(S' \vert S); (i, I_k))$ *is symmetric, then the unconditional proposal* $Q(S' \vert S)$ *is symmetric. In particular, the acceptance ratio of the inner M-H step 9 does not depend neither on* p_k *and nor on* q_μ.

The Bayesian estimator requires specifying prior distributions and proposal densities. All priors $p(\theta)$ are normal and all proposals (p_k, μ, and q_μ) are uniform over their respective domains.

4.3 Parametrization

The payoffs from (1) and (2) have six sets of parameters: v_i, w_{ij}, q, ϕ, ϕ_S and ϕ_N. In the empirical specification, the first three are functions of the data $v_i = V(X_i)$, $w_{ij} = W(X_i, X_j)$, $q_{ijk} = q(X_i, X_j, X_k)$. Careful scrutiny of the data and extensive experimentation with various parametrizations motivate the final specification which is discussed in the online appendix.

4.4 Identification

Because the model pertains to the exponential family, identification within the framework of many networks follows immediately. Indeed, a corollary of theorem 1 is that the likelihood of the model is proportional to $\exp\left\{\sum_{r=1}^{R}\theta_i w_i(S,X)\right\}$, where $w_i : \mathbf{S}\times\mathbf{X}\longrightarrow\mathbb{R}$ are functions of the data. To obtain identification, it is enough that the sufficient statistics w_i are linearly independent functions on $\mathbf{S}\times\mathbf{X}$ (e.g., see Lehmann and Casella (1998) for a textbook treatment). In the structural model above, this condition is readily established.[19]

Unobservable heterogeneity in friendship selection and decision to smoke

In addition to the models' parameters for observable attributes, it is possible to incorporate agents' specific unobservable types $\tau_i \sim N(0,\sigma_\tau^2)$ which may influence *both* the utility for friendships, e.g. $W(.,.)$ could include term $|\tau_i-\tau_j|$, and also the propensity to smoke, e.g. $V(.)$ could include a term $\rho_\tau\tau_i$. In this case the likelihood has to integrate out $\vec{\tau}$:

$$p(S|\theta)=\int_{\vec{\tau}}\frac{\exp\{\mathcal{P}_\theta(S,\vec{\tau}\}}{\sum_{\hat{S}}\exp\{\mathcal{P}_\theta(\hat{S},\vec{\tau})\}}\phi(\vec{\tau})d\vec{\tau} \tag{11}$$

There are a couple of approaches to discuss identification in this case. Within the Bayesian paradigm, identification casually obtains as long as the data provides information about the parameters. Even a weakly informative prior can introduce curvature into the posterior density surface that facilitates numerical maximization and the use of MCMC methods. However, the prior distribution is not updated in directions of the parameter space in which the likelihood function is flat (see An and Schorfheide, 2007). From a frequentist perspective, the heuristic identification argument goes as follows. Friends who are far away in observables, must have realizations of the unobservables very close by. If in the data those individuals are either smokers

[19] Most of the parameters are identified in the asymptotic frame where the size of the network grows to infinity (as opposed to the number of networks going to infinity). For example, turning off the externalities ($\phi=0$, $\phi_S=0$, $\phi_N=0$, $q=0$, $\psi=0$) implies that both smoking and friendships are independently distributed so that standard LLNs apply in the single large network asymptotics.

Table 2: Parameter estimates (posterior means)

	Utility of smoking				
	Parameter	No net data	Exog net	No PE	Model
1	Baseline probability of smoking	0.12^{***}	0.17^{***}	0.21^{***}	0.18^{***}
2	Price ×100	-0.17	-0.21	-0.61^{***}	-0.24^{*}
3	Mom edu (HS&CO)^{MP}	-0.04^{***}	-0.05^{***}	-0.05^{***}	-0.05^{***}
4	HH smokes	0.11^{***}	0.13^{***}	0.16^{***}	0.14^{***}
5	Grade 9+^{MP}	0.18^{***}	0.16^{***}	0.24^{***}	0.16^{***}
6	Blacks^{MP}	-0.3^{***}	-0.3^{***}	-0.35^{***}	-0.31^{***}
7	30% of the school smokes^{MP}	0.07^{***}	0.05^{***}	–	0.05^{***}

	Utility of friendships				
	Parameter	No net data	Exog net	No PE	Model
8	Baseline number of friends	–	–	4.63^{***}	3.4^{***}
9	Different sex$^{MP\%}$	–	–	-0.72^{***}	-0.72^{***}
10	Different grades$^{MP\%}$	–	–	-0.89^{***}	-0.89^{***}
11	Different race$^{MP\%}$	–	–	-0.33^{*}	-0.39^{***}
12	Cost/Economy of scale	–	–	-0.21^{***}	-0.22^{***}
13	Triangles$^{MP\%}$	–	–	1.18^{***}	1.22^{***}
14	ϕ^{MP}_{smoke}	–	0.04^{***}	–	0.05^{***}
15	$\phi^{MP}_{nosmoke}$	–	0.03^{***}	–	0.04^{***}

Note: MP stands for the estimated marginal probability in percentage points and MP% for estimated marginal probability in percent, relative to the baseline probability. The posterior sample contains 10^5 simulations before discarding the first 20%. The shortest 90/95/99% credible sets not including zero is indicated by */**/*** respectively.

or non smokers with very high probability then it must be the case that ρ_τ is large. However, formalizing this argument is nether immediate nor it is clear whether this argument will support non-parametric identification so this endeavor is left for future research.

4.5 Estimation results

Table 2 presents model's estimates (the posterior means) for four different estimation scenarios: (i.) without network data, (ii.) with fixed network, (iii.) without peer effects, and (iv.) the full model. The estimates have been transformed for ease of interpretation to baseline probabilities, marginal probabilities (MP in ppt) and relative marginal probabilities ($MP\%$ in pct).[20] It is worth pointing out that the estimate for

[20]Details about the parametrization and re-parametrization are available in the online appendix.

the price coefficient does not vary much in magnitude (but only in significance). The point estimates in table 2 together with the posterior distributions of this parameter in figure 3 suggest that the largest biases arise when peer effects terms are omitted (column "No PE") or when the econometrician does not have data on the friendship network (column "No Net Data").[21] Nevertheless, it is difficult to interpret the magnitudes of these differences nor the magnitudes of the structural estimates altogether in a concrete economic context. This is the case because the reported marginal effects are first order approximations which do not take into account the overall equilibrium response of the system.[22]

A final point on the estimation results is that the peer effect externalities are very different between smokers compared to those between non-smokers. Figure 4 reveals that the peer pressures between smokers is much stronger than that of non-smokers.(see footnote 22)

5 Policy experiments

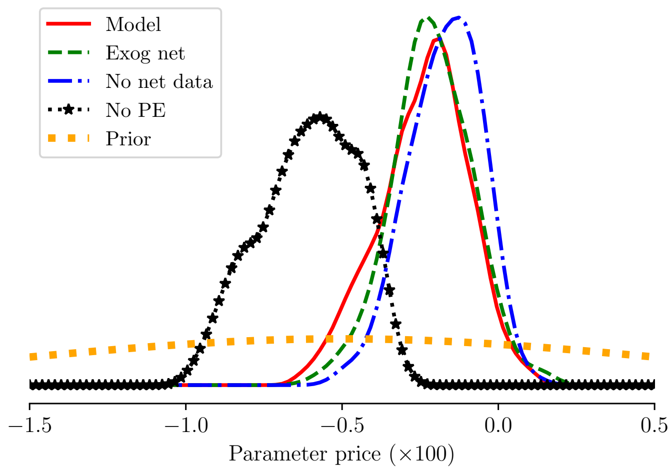
5.1 A. Changes in the price of tobacco

The estimated model serves as a numerical prototype for the equilibrium behaviors and, in particular, for the equilibrium adjustments to various policy interventions. Table 3 presents simulated increases in tobacco prices ranging from 20 to 160 cents (in the sample tobacco prices average at \$1.67 for a pack) and their effect on the overall tobacco smoking rates for the sample. The table compares the predictions from the full model to those from the model when agents are restricted from adjusting their friendship links and those from a model that is estimated without data on the friendship network.

[21] The hypotheses of equal means between the model's posterior and each of the other posteriors in figure 3 are rejected with $p < 0.01$ by t-tests.

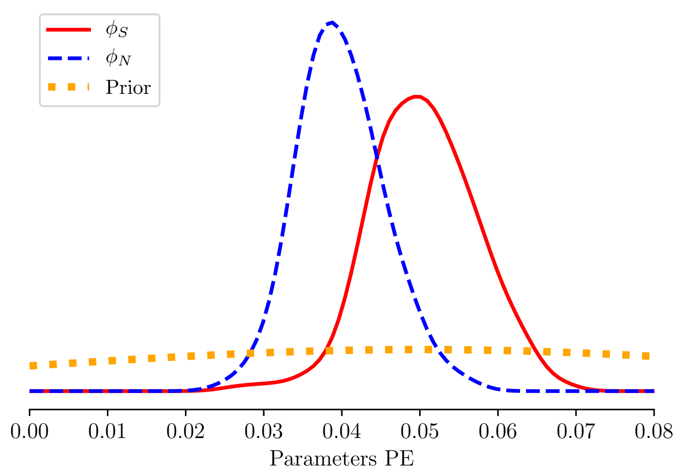
[22] A related point is that the parameter ϕ_S cannot be interpreted as the effect on the likelihood of smoking from a randomly assigned friend who is a smoker because, in the model, individuals cannot be forced into friendships. Rather, individual's utility increases with ϕ_S (or ϕ_N) with every instance where her choice to smoke (or not) and her choice of a friend are such that she and this friend of hers both smoke (or not).

Figure 3: Posterior distribution for the price parameter



Note: The hypotheses for equal means between the model's posterior and each of the other posteriors on the plot are rejected with $p < 0.01$ by t-tests.

Figure 4: Posterior distribution for the local PE parameters



Note: The hypotheses for equal means and equal distributions between the parameters for peer effects among smokers ϕ_S and among non-smokers ϕ_N are rejected with $p < 0.01$.

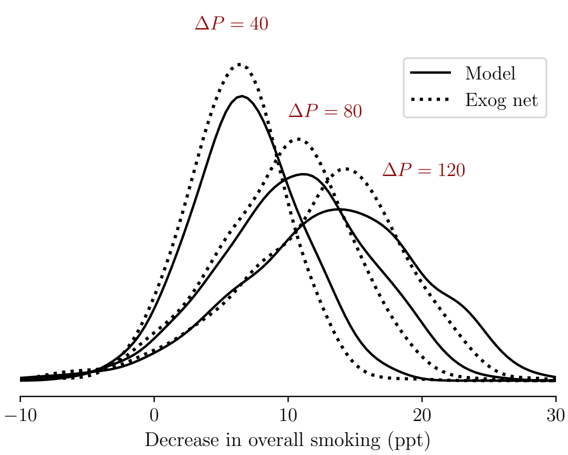
As seen in table 3, smoking rates responds to price changes. Comparison between model's predictions with and without friendship adjustments (columns two and three)

Table 3: The effect on smoking rate from changes in the price of tobacco

Price increase	Model	Exog net	No net data
20	2.5	2.2	1.3
40	4.7	4.2	2.6
60	6.9	6.1	3.9
80	8.7	7.9	5.1
100	10.3	9.4	6.2
120	11.8	10.9	7.4
140	13.1	12.3	8.4
160	14.3	13.5	9.5

Note: The first column shows proposed increases in tobacco prices in cents. The average price of a pack of cigarettes is \$1.67 so that 20 cents is approximately 10%. The second through fourth columns show the predicted increase in the overall smoking (baseline 41%) in ppt from the full model, from the model when the friendship network is fixed, and from the model when no social network data is available (i.e., $\phi_S = \phi_N = 0$).

Figure 5: Distribution of the effect on smoking rate from selected price changes



Note: In the model with exogenous friendships the distribution of the predicted smoking rates underestimates both the mean smoking rates (table 3) and the amount of uncertainty associated with the policy intervention (i.e, the variance of the distribution) as compared with the full model. The appendix provides statistical tests for these differences.

reveals that the latter underestimates the mean response by around 15% . In addition, the model without friendship choices underestimates the variance of this response as well (see figure 5). Finally, lack of network data (forcing the restriction $\phi_S = \phi_N = 0$) leads to a bias in the mean response to price changes that is between 50% and 70% of the prediction of the full model.

This analysis suggests that the freedom of breaking friendships and changing smoking behavior induces slightly larger decrease in overall smoking compared to situation when individuals are held in their existing (fixed) social networks. Figuratively, a price change has two effects on the decision to smoke: the direct effect operates through changing individuals' exogenous decision environment and the indirect/spillover effect operates through changing the peer norm which then puts additional pressure on the individuals' to follow the change. When comparing the endogenous to fixed network, the direct effect is likely to be stronger in the former environment while the indirect effect is likely to be stronger in the latter environment.[23] Related to this decomposition, this study suggests that quantitatively the direct effect dominates in shaping the overall equilibrium adjustments.

5.2 B. Changes in the racial composition of schools

Suppose that in a given neighborhood there are two racially segregated schools: "White School" consisting of only white students and "Black School" consisting of only black students. One would expect that the smoking prevalence in White school is much higher compared to Black school because, in the sample, black high students smoke three times less than white high school students. Consider a policy aiming to promote racial desegregation, which prevents schools from enrolling more than x percent of students of the same race. If such policy is in place, will students from different races form friendships and will these friendships systematically impact the

[23]It is interesting to relate this findings to the theoretical analysis in Jackson (2018) who argues that variability in individuals' popularity (degree in a social network) leads to biased perceptions for the social norm which in turn leads to higher levels of activities compared to a situation when there is no variability in individuals' popularity. This counterfactual experiments hints to such amplification mechanisms (in quite different settings).

Table 4: The effect on smoking rates from same-race students caps

Same-race cap (%)	School White	School Black	Overall
0	32.9	4.5	18.7
10	29.2	6.7	17.9
20	25.6	9.3	17.4
30	23.6	11.1	17.4
40	18.8	15.0	16.9
50	17.0	16.8	16.9

Note: A cap of $x\%$ same-race students is implemented with a swap of $(100 - x)\%$ students. The last column shows the predicted changes in overall smoking under different same-race caps. The policy induces statistically significant changes in the overall smoking as suggested by the statistical tests in the appendix.

overall smoking in one or another direction?

One of the racially balanced schools in the sample is used to evaluate the effect of this policy.[24] In particular, the Whites and the Blacks from this school serve as prototypes for the White School and Black School respectively. To implement the proposed policy, a random set of students from the White School is swapped with a random set of students from the Black School. For example to simulate the effect of a 70% cap on the same-race students in a school, a swap of 30% is simulated.

Table 4 presents the simulation results, which suggest that racial composition affects the overall smoking prevalence. The first column shows the size of the set of students which is being swapped. The second, third, and forth columns show the simulated smoking prevalence in the White School, Black School, and both, respectively. The table suggest that overall smoking prevalence is lower when schools are racially balanced, thus supporting policies promoting racial integration in the context of fighting high smoking rates.[25]

It is important to note that the simulations here offer only suggestive evidence on the role of racial desegregation on the overall prevalence of smoking. There are

[24]The school has 150 students of which 40% are Whites and 42% are Blacks. It incorporates students from grades 7 to 12. From these, the simulations use students from grades 10 to 12 because older students are more likely to form meaningful friendships and to smoke.

[25]The appendix demonstrates that these differences have statistical power.

many factors, e.g. the profile of all observables for the entire schools (income, home environment, tobacco price, etc), that are likely to influence the outcome of desegregation. Unfortunately, the Add Health data does not offer substantial variation in those factors and the empirical analysis relies on a (the only) racially balanced school in the data. The author hopes this study to stimulate further research into this question.

5.3 C. Aggregate effects of an anti-smoking campaign

The last experiment considers the effects of an anti-smoking campaign that can prevent with certainty a given number of students from smoking. An example of such intervention is a weekend-long information camp on the health consequences of smoking. Assuming that the camp is very effective in terms of preventing students from smoking but it is too costly to enroll all students in this camp, the question is once the "treated students" come back will their smoking friends follow their example and stop smoking, or will their friends un-friend them and continue smoking?

Table 5: Spillovers

Campaign (%)	Smoking	Predicted effect proportional	Actual effect	Multiplier
-	42.1	-	-	
3	39.6	1.3	2.6	2.0
5	38.2	2.1	3.9	1.9
10	34.6	4.2	7.5	1.8
20	28.7	8.4	13.4	1.6
30	23.5	12.6	18.6	1.5
50	15.1	21.1	27.0	1.3

Note: The first column lists the alternative attendance rates. The second and third columns display the smoking rate and the change in smoking rate respectively if the decrease would be proportional to the intervention, i.e. computes a baseline without peer effects. The last column computes the ratio between the percentage change in the number of smokers and the attendance rate. Note that that attendance is random with respect to the smoking status of the students. If the campaign is able to target only students who are currently smokers, the spillover effects will be even larger.

Table 5 presents the simulation results with two schools that feature smoking rates at the sample mean. The table suggests that an anti-smoking campaign may have a large impact on the overall prevalence of smoking, without necessarily being able to directly engage a large part of the student population.[26] In particular, the multiplier factor–the ratio between the actual effect and effect constrained to the treated sub-population–indicated a substantial spillover effects reaching up to the factor of 2. These spillover effects operate through the social network, from those who attended the camp to the rest of the school.

6 Concluding remarks

The premise of this paper is that individuals may respond differently to changes, with some following their friends' behaviors and others breaking away from their old friends in a search for new friends that will accept their new behaviors. This decision environment involves fundamentally different choices and generates complex mathematical structures. In equilibrium, players internalize the need for consensus in forming friendships and choose their optimal strategies on subsets of k players - a form of bounded rationality. The k-player consensual dynamic delivers a probabilistic ranking of the proposed equilibria, and, via a varying k, facilitates the implementation of the model.

The estimation of a structural model of adolescents' smoking and friendships demonstrates that peer effect complementarities between smokers are substantially stronger than those between non-smokers. It also documents the estimation biases due to not accounting for the endogeneity of the friendship network and those due to the lack of social network data. Counterfactual analysis with the estimated model suggests that: (a.) the response of the friendship network to changes in tobacco price amplifies the intended effect of price changes on smoking, (b.) racial desegregation of high-schools decreases the overall smoking prevalence, (c.) the peer effect complementarities are substantially stronger between smokers compared to between

[26]The policy is simulated 10^3 times, where each time a new random draw of attendees is being considered.

non-smokers, (d.) the magnitude of the spillover effects from small scale policies targeting individuals' smoking choices are roughly double compared to the scale of these policies.

Overall this paper formulates an avenue to study the complementarities and coordination in live social networks, i.e. social networks that adapt to the behaviors of individuals. The literature has just started to understand the forces present in these environments (e.g., see Jackson (2018)) while the empirical investigation of many hypothesis remains for the future (e.g., Carrell et al. (2013), Graham et al. (2014)).

A Proofs

Proof (Proposition 1(on p. 13)) Note that $\Delta_{g_{ij}} u_i() = \Delta_{g_{ij}} u_j()$. This property of the preferences implies that the unconstrained maximum in (4) is feasible w.r.t. the stability constraints (5). That is, for any i and $I_k = \{i\} \cup \{i_1, \ldots, i_{k-1}\}$ the solution of individual's decision problem (4-5) is simply

$$\operatorname*{argmax}_{\substack{a_i, g_{ij} \\ j \in I_k \setminus i}} \mathcal{P}(S). \tag{4}$$

This completes the proof of part one because (4) always has a solution.

For part two, the first step is to extend the property $\Delta_{g_{ij}} u_i() = \Delta_{g_{ij}} u_j()$ to a deeper property of the preferences namely that the preferences of all players can be expressed by a single potential function.[27] Indeed, consider $\mathcal{P} : \mathbf{S} \times \mathbf{X}_n \longrightarrow \mathbb{R}$:

$$\begin{aligned}
\mathcal{P}(S, X) &= \sum_i a_i v(X_i) + \frac{1}{2} \sum_{i,j} g_{ij} w(X_i, X_j) && (12) \\
&\quad + \frac{1}{2} \phi \sum_{i,j; i \neq j} a_i a_j + \frac{1}{2} \phi_S \sum_{i,j} g_{ij} a_i a_j + \frac{1}{2} \phi_N \sum_{i,j} g_{ij} (1 - a_i)(1 - a_j) && (13) \\
&\quad + \frac{1}{6} \sum_{i,j,k} q(X_i, X_j, X_k) g_{ij} g_{jk} g_{ki} && (14)
\end{aligned}$$

where $i \neq j$ is dropped from the summation ranges where possible because the convention that g_{ii} is defined and equals to 0 for all i so that $\sum_{i,j; i \neq j} g_{ij} = \sum_{i,j} g_{ij}$. To show that $\mathcal{P}$ is potential, it is sufficient to verify that (using assumption 1):

$$\begin{aligned}
\Delta_{a_i} u_i() &= v_i + \phi \sum_{j \neq i} a_j + \phi_S \sum_{j \neq i} g_{ij} a_j - \phi_N \sum_{j \neq i} g_{ij} (1 - a_j) && (15) \\
&= \Delta_{a_i} \mathcal{P}() \\
\Delta_{g_{ij}} u_i() &= w_{ij} + q(X_i, X_j, X_k) \sum_k g_{ik} g_{jk} && (16) \\
&\quad - \psi(d_i + d_j) + \phi_S a_i a_j + \phi_N (1 - a_i)(1 - a_j) && (17) \\
&= \Delta_{g_{ij}} \mathcal{P}()
\end{aligned}$$

[27] The existence of potential implies $\Delta_{g_{ij}} u_i() = \Delta_{g_{ij}} u_j()$ but the converse is not true.

Next, fix k and consider the following adaptive dynamic on $\mathbf{S}$. Every period draw at random i and I_k (from the uniform distributions over their respective domains), and let i choose in her argmax (4). For this dynamic, the value of the potential is nondecreasing so, invoking submartingale convergence argument, the potential convergences. Unless two states have the same potential (generically false), this implies that the state converges to a particular network which is, of course, a NEkS network. This same technology appears in the proof of proposition 3. ■

As it will be useful later on, proposition 5 states characterization (4) in both directions. The proof of the if direction follows closely that of the only if direction above, and is omitted.

Proposition 5 *S^* is a NEkS network* iff $\forall i, I_k = \{i\} \cup \{i_1, \ldots, i_{k-1}\}$

$$(a_i^*, g_{ij}^*)_{j \in I_k \setminus i} \in \underset{\substack{a_i, g_{ij} \\ j \in I_k \setminus i}}{\operatorname{argmax}} \mathcal{P}((a_i, g_{ij})_{j \in I_k \setminus i}; S^*_{-(a_i, g_{ij})_{j \in I_k \setminus i}})$$

PROOF (PROPOSITION 2(ON P. 13)) For $k = 2$, definition 1 directly implies that a NEkS network is pairwise stable. Note that this observation is independent of the particular payoff structure here.

Let $k = n$. That a NEkS network is pairwise stable follows from part 3 of this proposition (demonstrated next). To see that a NEkS network S^* is a Nash network, consider the following strategies in a normal form link-announcement game (given the equilibrium behavior $\vec{a}^*$): each player announces his NEkS links. Proceeding by contradiction, for if a player has a profitable deviation then it would be possible to construct (appending a_i^*) an $S_{(i)}$ which she prefers to her NEkS play $S^*_{(i)}$. Therefore $S^*_{(i)} \notin \operatorname{argmax}_{\substack{a_i, g_{ij} \\ j \in I_k \setminus i}} u_i(S) = \operatorname{argmax}_{\substack{a_i, g_{ij} \\ j \in I_k \setminus i}} \mathcal{P}(S)$ which contradicts proposition 5.

Finally, the characterization from proposition 5 directly implies part three. In particular, if $k' < k$, $I_{k'} \subset I_k$ and $(a_i^*, g_{ij}^*)_{j \in I_k \setminus i} \in \operatorname{argmax}_{\substack{a_i, g_{ij} \\ j \in I_k \setminus i}} \mathcal{P}((a_i, g_{ij})_{j \in I_k \setminus i}; S^*_{-(a_i, g_{ij})_{j \in I_k \setminus i}})$ then $(a_i^*, g_{ij}^*)_{j \in I_{k'}} \in \operatorname{argmax}_{\substack{a_i, g_{ij} \\ j \in I_{k'}}} \mathcal{P}((a_i, g_{ij})_{j \in I_{k'}}; S^*_{-(a_i, g_{ij})_{j \in I_{k'}}})$. ■

PROOF (PROPOSITION 3 (ON P. 15) That any NEkS network is absorbing for the kCD follows from definition 1. The second part follows from observing that $\mathcal{P}_t$ is a submartingale, i.e., $E[\mathcal{P}_{t+1}|S_t] \geq \mathcal{P}_t$., so that $\{\mathcal{P}_t\}$ converges almost surely. Because the network size is finite it follows that $\{\mathcal{P}_t\}$ is constant for large t and, generically, the same holds for S_t, i.e. $S_t = S^*$ for large enough t. Because of assumption 2 (any meeting is possible), this can happen only if S^* is a NEkS network. ■

Proof (Theorem 1 (p. 16)) The first part follows from standard results on convergence of Markov chains. In particular, k-CDs with random utility induce a finite state Markov chain which, with assumptions 2 and 3, is irreducible, positive recurrent, and aperiodic. This is sufficient to obtain the conclusion of part one.

For the second part, it is enough to show that

$$\Pr(S'|S;k)\exp\{\mathcal{P}(S)\} = \Pr(S|S';k)\exp\{\mathcal{P}(S')\}, \tag{18}$$

where $\Pr(S'|S;k)$ is the one step transition probability for moving from S to S'.

There are two cases to consider: $\Pr(S'|S;k) = 0$ and $\Pr(S'|S;k) > 0$. Note that the hypothesis guarantees that $\Pr(S', S;k) > 0$ iff $\Pr(S, S';k) > 0$. Thus, if $\Pr(S'|S;k) = 0$ then $\Pr(S|S';k) = 0$ and, trivially, (18) holds.

Consider the case $\Pr(S'|S;k) > 0$. For fixed k, S, and S let $\mathbf{M}_{S'|S;k}$ be the set of all possible meetings that can result in transitioning from S to S'. Note that for some triples (S, S', k), $\mathbf{M}_{S'|S;k}$ is empty. However, if $\Pr(S'|S;k) > 0$ then $\mathbf{M}_{S'|S;k} \neq \emptyset$.

Let us pause with an example of this notation. Given the triple (S, S', k)

$$\Pr(S'|S;k) = \sum_{\mu \in \mathbf{M}_{S'|S;k}} \Pr(\mu) \frac{\exp\{u_i(S')\}}{\sum_{\hat{S} \in \mathbf{N}_k(\mu,S)} \exp\{u_i(\hat{S})\}}.$$

Consider the case when S and S' agree on all $\{g_{ij}\}_{i \neq j}$ but differ in a_i for some i, say $S = (a_i = 0, S_{-a_i})$ and $S' = (a'_i = 1, S_{-a_i})$. Then, $\mathbf{M}_{S'|S;k}$ is the set of all possible meeting tuples (i, I_{k-1}) where player i meets different $\{i_1, \ldots, i_{k-1}\}$, and the size of $\mathbf{M}_{S'|S;k}$ is $\binom{n-1}{k-1}$. To close the example, assume that all meetings are equally likely and that individuals are indifferent to all outcomes (i.e. u_i is a constant). Then

$\Pr(\mu) = \frac{1}{n}\frac{1}{\binom{n-1}{k-1}}$ and $\frac{\exp\{u_i(S')\}}{\sum_{\hat{S}\in\mathbf{N}_k(\mu,S)}\exp\{u_i(\hat{S})\}} = \frac{1}{2^k}$ so that

$$\Pr(S'|S;k) = \binom{n-1}{k-1}\frac{1}{n\binom{n-1}{k-1}}\frac{1}{2^k} = \frac{1}{n2^k}.$$

Recall that $\mathbf{N}_k(S,\mu) \subset \mathbf{S}_n$ denotes the set of all possible states that can result from the meeting μ following a state S. The proof follows from the following observations:[28]

Lemma 1 *For all k, S, S', and $\mu = (i, I_{k-1})$:*
(i) $\mathbf{M}_{S'|S;k} = \mathbf{M}_{S|S';k}$ *for all* $S, S' \in \mathbf{S}_n$*;*
(ii) $S' \in \mathbf{N}_k(\mu, S)$ *iff* $S \in \mathbf{N}_k(\mu, S')$*;*
(iii) If $S' \in \mathbf{N}_k(\mu, S)$ *then* $\mathbf{N}_k(\mu, S) = \mathbf{N}_k(\mu, S')$*.*

Part (i) asserts that each meeting that can result in transitioning from S to S' may result in transitioning from S' to S as well (provided the starting state were S'). Part (ii) re-states this observation in terms of the neighborhoods of S and S' given a meeting μ. Finally, part (iii) notes that if a meeting μ could result in S transiting to S', then the set of all feasible states following μ and S coincides with the set of all feasible states following μ and S'.

From lemma 1, the one step transition probability can be written as:

$$\begin{aligned}
\mathcal{P}(S)\Pr(S'|S;k) &= \mathcal{P}(S)\sum_{\mu\in\mathbf{M}_{S'|S;k}}\Pr(\mu)\frac{\exp\{u_i(S')\}}{\sum_{\hat{S}\in\mathbf{N}_k(\mu,S)}\exp\{u_i(\hat{S})\}} && (19)\\
&= \mathcal{P}(S)\sum_{\mu\in\mathbf{M}_{S|S';k}}\Pr(\mu)\frac{\exp\{\mathcal{P}(S')\}}{\sum_{\hat{S}\in\mathbf{N}_k(\mu,S)}\exp\{\mathcal{P}(\hat{S})\}} && (20)\\
&= \mathcal{P}(S')\sum_{\mu\in\mathbf{M}_{S|S';k}}\Pr(\mu)\frac{\exp\{\mathcal{P}(S)\}}{\sum_{\hat{S}\in\mathbf{N}_k(\mu,S')}\exp\{\mathcal{P}(\hat{S})\}} && (21)\\
&= \mathcal{P}(S)\Pr(S,S';k) && (22)
\end{aligned}$$

Where the particular expression for $Pr(S'|S,\mu) = \frac{\exp\{u_i(S')\}}{\sum_{\hat{S}\in\mathbf{N}_k(\mu,S)}\exp\{u_i(\hat{S})\}}$ follows from assumption 4 on the distribution of the error term. ∎

[28] The proof of lemma 1 involves basic reasoning and is omitted. The challenging part is to state and interpret the lemma. Formal proof available upon request.

Proof (Theorem 2 (p. 17)) Because there is no natural ordering of $\mathbf{S}_n$, use functions as opposed to vectors in the eigenproblem. For $I \subset \{(i,j) : i \geq j\}$, define $\mathrm{e}_I : \mathbf{S}_n \to \mathbb{R}$ as

$$\mathrm{e}_I(S) = \prod_{i \neq j \in I} (-1)^{g_{ij}} \prod_{i = j \in I} (-1)^{a_{ij}} \tag{23}$$

with $\mathrm{e}_\emptyset(S) = 1$ for all S. Next, define

$$\lambda_{k,I} = \frac{\sum_{i \in \{i:(i,i) \notin I\}} \binom{n-1-|I_i|}{k-1}}{n\binom{n-1}{k-1}} \tag{24}$$

where $I_i = \{j : (i,j) \in I, i \neq j\}$

Lemma 2 *There are $2^{n(n+1)/2}$ pairs of $(\lambda_{k,I}, \mathrm{e}_{k,I})$ such that*

(i) $\sum_S \mathrm{e}_{k,I}(S)\mathrm{e}_{k,I'}(S) = 0$ if $I \neq I'$ and $\sum_S \mathrm{e}_{k,I}(S)\mathrm{e}_{k,I}(S) = 2^{n(n+1)/2}$

(ii) For any $S \in \mathbf{S}_n$

$$\sum_{S'} \Pr(S'|S;k)\mathrm{e}_I(S') = \lambda_{k,I}\mathrm{e}_I(S). \tag{25}$$

The first part of the lemma is trivial to verify. For the second part, write:

$$\sum \Pr(S'|S)\mathrm{e}_I(S') = \sum_{S'}\sum_{\mu} \Pr(\mu)\Pr(S'|S,\mu)\mathrm{e}_I(S') \tag{26}$$

$$= \sum_{S'}\sum_{\mu \in \{\mu \cap I = \emptyset\}} \Pr(\mu)\Pr(S'|S,\mu)\mathrm{e}_I(S') + \tag{27}$$

$$+ \sum_{S'}\sum_{\mu \in \{\mu \cap I \neq \emptyset\}} \Pr(\mu)\Pr(S'|S,\mu)\mathrm{e}_I(S') \tag{28}$$

$$= \sum_{S'}\sum_{\mu \in \{\mu \cap I = \emptyset\}} \Pr(\mu)\Pr(S'|S,\mu)\mathrm{e}_I(S') \tag{29}$$

Terms (28) vanish because whenever $\mu \cap I \neq \emptyset$ then $\sum_{S' \in \mathbf{N}_k(S,\mu)} \Pr(S'|S,\mu)\mathrm{e}_I(S') = 0$, as this summation involves 2^k terms and for half of these terms $\mathrm{e}_I(S') = \mathrm{e}_I(S)$ while for the other half $\mathrm{e}_I(S') = -\mathrm{e}_I(S)$, implying that $\sum_{\mu \in \{\mu \cap I \neq \emptyset\}} \sum_{S' \in \mathbf{N}_k(S,\mu)} \Pr(S'|S,\mu)\mathrm{e}_I(S')$ equals to 0.

Finally, note that if $\mu \in \{\mu \cap I = \emptyset\}$, i.e. $\mu = \{(i,i),(i,i_1),\ldots(i,i_{k-1})\} \cap I = \emptyset$

then for any $S' \in \mathbf{N}_k(S,\mu)$ we have that $\mathrm{e}_I(S) = \mathrm{e}_I(S')$ so that for (29) we can write

$$\begin{aligned}
\sum \Pr(S'|S)\mathrm{e}_I(S') &= \mathrm{e}_I(S)\Pr(\mu)\sum_{S'}\sum_{\mu\in\{\mu\cap I=\emptyset\}}\Pr(S'|S,\mu) &(30)\\
&= \mathrm{e}_I(S)\frac{1}{n\binom{n-1}{k-1}}\sum_{\mu\in\{\mu\cap I=\emptyset\}}\sum_{S'}\Pr(S'|S,\mu) &(31)\\
&= \mathrm{e}_I(S)\frac{1}{n\binom{n-1}{k-1}}\sum_{i\in\{i:(i,i)\notin I\}}\binom{n-1-|I_i|}{k-1} &(32)
\end{aligned}$$

because, by assumption, $\Pr(\mu) = \frac{1}{n\binom{n-1}{k-1}}$ and $\sum_{S'}\Pr(S'|S,\mu) = 1$. This completes the proof of lemma 2. To complete the proof of the theorem note that $\lambda_{k,I}$ are decreasing in $|I|$, so that the (second) largest $\lambda_{k,I}$ is achieved when $I = \{(i,j)\}$ with $i \neq j$. ■

PROOF (THEOREM 3 (P. 19) The proof follows immediately from the expression for the stationary distribution obtained in theorem 1 and proposition 5. ■

PROOF (PROPOSITION 4 (P. 21) For fixed $S, S' \in \mathbf{S}_n$ let $\mathbf{K}_{S'|S} \subset \{2,3,\dots,n\}$ be the set of all possible meeting sizes consistent with transition from S to S' of the k-PD. Recall that, for fixed k, $\mathbf{M}_{S'|S;k}$ is the set of all possible meetings that may induce transitioning from S to S'. The argument bellow follows from lemma 1, together with the observation that $\mathbf{K}_{S'|S} = \mathbf{K}_{S|S'}$. Indeed, the unconditional proposal Q from the algorithm in table 1 can be written as:

$$\begin{aligned}
Q(S'|S) &= \sum_{k\in\mathbf{K}_{S'|S}} p_k(k)\sum_{\mu\in\mathbf{M}_{S'|S}}\Pr(\mu)\frac{1}{|\mathbf{N}_k(\mu,S)|} &(33)\\
&= \sum_{k\in\mathbf{K}_{S|S'}} p_k(k)\sum_{\mu\in\mathbf{M}_{S|S'}}\Pr(\mu)\frac{1}{|\mathbf{N}_k(\mu,S)|} &(34)\\
&= \sum_{k\in\mathbf{K}_{S|S'}} p_k(k)\sum_{\mu\in\mathbf{M}_{S|S'}}\Pr(\mu)\frac{1}{|\mathbf{N}_k(\mu,S')|} &(35)\\
&= Q(S|S')
\end{aligned}$$

■

References

Ali, Mir M and Debra S Dwyer, "Estimating peer effects in adolescent smoking behavior: a longitudinal analysis.," *Journal of Adolescent Health*, 2009, *45* (4), 402–408.

_ and _ , "Estimating peer effects in adolescent smoking behavior: a longitudinal analysis," *Journal of Adolescent Health*, 2009, *45* (4), 402–408.

An, Sungbae and Frank Schorfheide, "Bayesian analysis of DSGE models," *Econometric reviews*, 2007, *26* (2-4), 113–172.

Badev, Anton, "Discrete games in endogenous networks: Theory and policy," PhD Dissertation, University of Pennsylvania 2013.

Baetz, Oliver, "Social activity and network formation," *Theoretical Economics*, 2015, *10* (2).

Ballester, Coralio, Antoni Calvó-Armengol, and Yves Zenou, "Who's Who in Networks. Wanted: The Key Player," *Econometrica*, 09 2006, *74* (5), 1403–1417.

Battaglini, Marco, Eleonora Patacchini, and Edoardo Rainone, "Endogenous Social Connections in Legislatures," Working Paper 25988, National Bureau of Economic Research June 2019.

Bhamidi, Shankar, Guy Bresler, and Allan Sly, "Mixing time of exponential random graphs," *Ann. Appl. Probab.*, 12 2011, *21* (6), 2146–2170.

Bloch, Francis and Matthew Jackson, "The formation of networks with transfers among players," *Journal of Economic Theory*, 2007, *133* (1), 83–110.

_ and Matthew O. Jackson, "Definitions of equilibrium in network formation games," *International Journal of Game Theory*, Oct 2006, *34* (3), 305–318.

Blume, Lawrence E., "The Statistical Mechanics of Strategic Interaction," *Games and Economic Behavior*, 1993, *5*, 387–424.

— , **William A. Brock, Steven N. Durlauf, and Rajshri Jayaraman**, "Linear Social Interactions Models," *Journal of Political Economy*, 2015, *123* (2), 444–496.

Boucher, Vincent, "Conformism and self-selection in social networks," *Journal of Public Economics*, 2016, *136*, 30 – 44.

— , **Chih-Sheng Hsieh, and Lung Fei Lee**, "Specification and Estimation of Network Formation and Network Interaction Models with the Exponential Probability Distribution," Technical Report, SSRN Working Paper 2019.

Bourlés, Renaud, Yann Bramoullé, and Eduardo Perez-Richet, "Altruism in Networks," *Econometrica*, 2017, *85* (2), 675–689.

Bramoullé, Yann and Rachel Kranton, "Games Played on Networks," in "The Oxford Handbook of the Economics of Networks," Oxford University Press, 06 2016.

— , **Andrea Galeotti, and Brian Rogers**, *The Oxford handbook of the economics of networks*, Oxford University Press, 2016.

— , **Rachel Kranton, and Martin D'Amours**, "Strategic Interaction and Networks," *American Economic Review*, March 2014, *104* (3), 898–930.

Brock, William A. and Steven N. Durlauf, "Discrete Choice with Social Interactions," *The Review of Economic Studies*, 2001, *68* (2), 235–260.

— **and** — , "Identification of binary choice models with social interactions," *Journal of Econometrics*, September 2007, *140* (1), 52–75.

Cabrales, Antonio, Antoni Calvó-Armengol, and Yves Zenou, "Social interactions and spillovers," *Games and Economic Behavior*, 2011, *72* (2), 339–360.

Calvó-Armengol, Antoni, "Job contact networks," *Journal of Economic Theory*, 2004, *115* (1), 191 – 206.

— , **Eleonora Patacchini, and Yves Zenou**, "Peer Effects and Social Networks in Education," *Review of Economic Studies*, 2009, *76* (4), 1239–1267.

Canen, Nathan, Francesco Trebbi, and Matthew Jackson, "Endogenous Network Formation in Congress," Working Paper 22756, National Bureau of Economic Research October 2016.

Carrell, Scott E., Bruce I. Sacerdote, and James E. West, "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation," *Econometrica*, 05 2013, *81* (3), 855–882.

CDC, "Reducing Tobacco Use," Technical Report, Centers for Disease Control and Prevention 2000.

Chaloupka, Frank and Henry Wechsler, "Price, tobacco control policies and smoking among young adults," *Journal of Health Economics*, 1997, *16* (3), 359–373.

Chandrasekhar, Arun, "Econometrics of network formation," in "The Oxford Handbook of the Economics of Networks" 2015, pp. 303–357.

_ **and Matthew O. Jackson**, "A network formation model based on subgraphs," *Stanford Working Paper*, 2016.

Cournot, Augustin, *Researches into the Mathematical Principles of the Theory of Wealth*, The Macmillan Company, 1838.

de Paula, Áureo, "Econometrics of network models," Technical Report 16 2016.

_ **, Seth Richards-Shubik, and Elie Tamer**, "Identifying Preferences in Networks With Bounded Degree," *Econometrica*, 2018, *86* (1), 263–288.

Debreu, Gerard and I. N. Herstein, "Nonnegative Square Matrices," *Econometrica*, 1953, *21* (4), 597–607.

Foster, Dean and Peyton Young, "Stochastic evolutionary game dynamics," *Theoret. Population Biol*, 1990.

Frank, Ove and David Strauss, "Markov graphs," *Journal of the american Statistical association*, 1986, *81* (395), 832–842.

Goldsmith-Pinkham, Paul and Guido W Imbens, "Social networks and the identification of peer effects," *Journal of Business & Economic Statistics*, 2013, *31* (3), 253–264.

Goyal, Sanjeev and Fernando Vega-Redondo, "Network formation and social coordination," *Games and Economic Behavior*, February 2005, *50* (2), 178–207.

_ **and Sumit Joshi**, "Unequal connections," *International Journal of Game Theory*, 2006, *34* (3), 319–349.

Graham, Bryan S, "An econometric model of network formation with degree heterogeneity," *Econometrica*, 2017, *85* (4), 1033–1063.

_ **, Guido W Imbens, and Geert Ridder**, "Complementarity and aggregate implications of assortative matching: A nonparametric analysis," *Quantitative Economics*, 2014, *5* (1), 29–66.

Hiller, Timo, "Peer effects in endogenous networks," *Games and Economic Behavior*, 2017, *105* (C), 349–367.

Hsieh, Chih-Sheng and Lung Fei Lee, "A social interactions model with endogenous friendship formation and selectivity," *Journal of Applied Econometrics*, 2016.

_ **, Michael D König, and Xiaodong Liu**, "Network Formation with Local Complements and Global Substitutes: The Case of R&D Networks," Technical Report, University of Zurich 2016.

Jackson, Matthew O, "A survey of network formation models: stability and efficiency," *Group formation in economics: Networks, clubs, and coalitions*, 2005, *664*, 11–49.

Jackson, Matthew O., *Social and Economic Networks*, Princeton, NJ, USA: Princeton University Press, 2008.

_ , "The Friendship Paradox and Systematic Biases in Perceptions and Social Norms," *Journal of Political Economy*, 2018.

_ **and Alison Watts**, "The Existence of Pairwise Stable Networks," *Seoul Journal of Economics*, 2001, *14* (3).

_ **and** _ , "The Evolution of Social and Economic Networks," *Journal of Economic Theory*, October 2002, *106* (2), 265–295.

_ **and Asher Wolinsky**, "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 1996, *71* (0108), 44–74.

_ **and Brian W. Rogers**, "Meeting Strangers and Friends of Friends: How Random Are Social Networks?," *American Economic Review*, June 2007, *97* (3), 890–915.

_ **and Yves Zenou**, "Games on Networks," in "Handbook of Game Theory with Economic Applications," Vol. 4, Elsevier, 2015, chapter 3, pp. 95–163.

Johnsson, Ida and Hyungsik Roger Moon, "Estimation of Peer Effects in Endogenous Social Networks: Control Function Approach," 2017.

Kandori, Michihiro, George J. Mailath, and Rafael Rob, "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, January 1993, *61* (1), 29–56.

König, Michael D, Claudio J Tessone, and Yves Zenou, "Nestedness in networks: A theoretical model and some applications," *Theoretical Economics*, 2014, *9* (3), 695–752.

Krauth, Brian V, "Peer effects and selection effects on smoking among Canadian youth," *Canadian Journal of EconomicsRevue canadienne dconomique*, 2005, *38* (3), 735–757.

Lagerås, Andreas and David Seim, "Strategic complementarities, network games and endogenous network formation," *International Journal of Game Theory*, 2016, *45* (3), 497–509.

Lehmann, E L and George Casella, *Theory of Point Estimation*, Vol. 31, Springer, 1998.

Leung, Michael, "A random-field approach to inference in large models of network formation," 2014.

Liang, Faming, "A double Metropolis–Hastings sampler for spatial models with intractable normalizing constants," *Journal of Statistical Computation and Simulation*, 2010, *80* (9), 1007–1022.

Liu, Xiaodong, Eleonora Patacchini, and Yves Zenou, "Endogenous peer effects: local aggregate or local average?," *Journal of Economic Behavior and Organization*, 2014, *103*, 39 – 59.

Marschak, Jacob, "Binary choice constraints on random utility indications," in Kenneth Arrow, ed., *Stanford Symposium on Mathematical Methods in the Social Sciences*, 1960, pp. 312–329.

McFadden, D, "Conditional logit analysis of qualitative choice behavior," *Frontiers in Econometrics*, 1974, *1* (2), 105–142.

Mele, Angelo, "A structural model of dense network formation," *Econometrica*, 2017, *85* (3), 825–850.

Menzel, Konrad, "Strategic network formation with many agents," Technical Report, Technical report, New York University 2015.

Monderer, Dov and Lloid S. Shapley, "Potential Games," *Games and Economic Behavior*, 1996, *14*, 124–143.

Murray, Iain, Zoubin Ghahramani, and David MacKay, "MCMC for doubly-intractable distributions," *Uncertainty in Artificial Intelligence*, 2006.

Myerson, Roger B., *Game theory - Analysis of Conflict*, Harvard University Press, 1991.

Nakajima, Ryo, "Measuring Peer Effects on Youth Smoking Behavior," *The Review of Economic Studies*, 2007, *74*, 897–935.

Nash, John, "Non-cooperative Games," PhD Dissertation, Princeton University 1950.

Powell, Lisa M, John A Tauras, and Hana Ross, "The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior.," *Journal of Health Economics*, 2005, *24* (5), 950–968.

Sheng, Shuyang, "A structural econometric analysis of network formation games," *USC-mimeo*, 2016.

Thurstone, L L, "A law of comparative judgment," *Psychological Review*, 1927, *34* (4), 273–286.

Train, Kenneth, *Discrete Choice Methods with Simulation*, SUNY-Oswego, Department of Economics, 2003.

Wasserman, Stanley and Philippa Pattison, "Logit models and logistic regressions for social networks: I. An introduction to Markov graphs andp," *Psychometrika*, 1996, *61* (3), 401–425.