

## **Linear Social Network Models**

**Lawrence E. Blume  
Cornell University  
IHS Vienna  
Santa Fe Institute  
lb19@cornell.edu**

**William A. Brock  
University of Wisconsin  
wbrock@ssc.wisc.edu**

**Steven N. Durlauf  
University of Wisconsin  
sdurlauf@ssc.wisc.edu**

**Rajshri Jayaraman  
ESMT, Berlin  
jayaraman@esmt.org**

**August 19, 2011**

**First Draft: Please do not circulate.**

Financial support has been supplied to Blume by NSF grant CCF-0910940 and by WWTF Grant “Die Evolution von Normen und Konventionen in der Wirtschaft”, Brock and Durlauf by the Vilas Professorship, and Durlauf by the University of Wisconsin Graduate School and Laurits Christensen Chair in Economics, all of which is greatly appreciated. Hon Ho Kwok and Xiangrong Yu have provided superb research assistance. We are grateful for comments from Youcef Msaid, Alex Rees-Jones, Dean Robinson, Michael Strain and Nichole Szembrot and to Charles Manski and Hashem Pesaran for discussions of the issues we address. This paper was written in honor of James J. Heckman, whose influence will be evident throughout.

Corresponding Author: Steven N. Durlauf, Department of Economics, University of Wisconsin, Madison WI 53706.

## Abstract

This paper provides a systematic analysis of identification in linear social networks models. This is both a theoretical and an econometric exercise in that it links identification analysis to a rigorously delineated model of interdependent decisions. We develop a Bayes-Nash equilibrium analysis for interdependent decisions under incomplete information in networks that produces linear strategy profiles of the type conventionally used in empirical work and which nests linear social interactions models as a special case. We consider identification of both contextual and endogenous social effects under alternative assumptions on the *a priori* information on network structure available to an analyst and contrast the informational content of individual-level and aggregated data. This analysis is then extended to an example of a two stage game in which networks form in the first stage and outcomes occur in the second. The effects of endogenous network formation on identification are then analyzed.

JEL Codes: C21, C23, C31, C35, C72, Z13

Keywords: social networks, identification, incomplete information games

...friendship...is...most necessary with a view to living. For without friends, no one would choose to live, though he had all other goods; even rich men and those in possession of office and of dominating power are thought to need friends most of all; for what is the use of such prosperity without the opportunity of beneficence, which is exercised chiefly and in its most laudable form towards friends? Or how can prosperity be guarded and preserved without friends? The greater it is, the more exposed is it to risk. And in poverty and in other misfortunes men think friends their only refuge. It helps the young too, to keep from error; it aids older people by ministering to their needs and supplementing the activities that fail from weakness; those in the prime of life it stimulates to noble actions — ‘two going together’ — for with friends men are more able to think and to act.

Aristotle, *Nicomachean Ethics*, 8.1<sup>1</sup>

## 1. Introduction

The study of social influences has become a major area of economic theory, econometrics, and empirical work, as evidenced by the surveys in Benhabib, Bisin, and Jackson (2010). Standard examples range across such disparate areas as the diffusion of technology (Conley and Udry (2010), Munshi (2004), Bandiera and Rasul (2006)), disease exposure (Miguel and Kremer (2004)), contraceptive practice (Kohler, Behrman, and Watkins (2001), Iyer and Weeks (2009)), smoking (Krauth (2006), Soetevent and Kooreman (2007), Nakajima (2007)), crime (Sirakaya (2006), Ballester, Calvó-Armengol, and Zenou (2010)), education (Cooley (2008), Bobonis and Finan (2009), Calvó-Armengol, Patacchini, and Zenou (2009), De Giorgio, Pellizzari, and Redaelli (2010)), the take up of public welfare programs (Bertrand, Luttmer and Mullainathan (2000), Aizer and Currie (2004)), labor market outcomes (Topa (2001), Munshi (2003), Bayer, Ross and Topa (2008)) and even obesity (Christakis and Fowler (2007), but see Cohen-Cole and Fletcher (2008)). This work is now broad enough to justify the claim that it constitutes a new field of “social economics,” a term that was

---

<sup>1</sup>Taken from *The Complete Works of Aristotle*, volume 2, J. Barnes ed., Princeton University Press, 6<sup>th</sup> printing, 1984.

proposed by Becker and Murphy (2001) when the consideration of social factors first began to play a major role in economic analyses.

Within the broad area of social economics, the study of social networks has arguably become the most prominent area of research. From the perspective of economic theory, social network analysis is now a well established area of specialization; see Jackson (2008) and Goyal (2009) as well as the relevant chapters in Benhabib, Bisin, and Jackson (2010) for overviews of the existing theory. Further, it is common for empirical papers to invoke social networks as the underlying structure by which social influences are transmitted; one standard example is information transmission.

While the theoretical and empirical literatures on social networks have expanded greatly over the last decade, there has been little contact between them. There has been remarkably little work on formal econometric issues pertaining to identification. This is especially surprising since there is now a rich literature on identification problems in social interactions models—a subset of social network models which assumes that individuals belong to predefined groups wherein all group members influence each other with equal intensity.<sup>2</sup>

The identification literature on social interactions has, since Manski (1993), demonstrated that important limits exist to identification. In particular, the so-called linear-in-means model, the workhorse of empirical research on social interactions, raises classical simultaneity problems.<sup>3</sup> This issue was first exposed by Manski, who dubbed it “the reflection problem”. The main exceptions to this absence of formal identification work for social networks models are Bramoullé, Djebbari, and Fortin (2009) and Blume, Brock, Durlauf, and Ioannides (2010).<sup>4</sup> Bramoullé, Djebbari and

---

<sup>2</sup>See Blume, Brock, Durlauf, and Ioannides (2010) for a review.

<sup>3</sup>“Linear-in-means” captures the idea that an individual’s behavior depends on the average behavior and/or characteristics of members of his group.

<sup>4</sup>Mention should also be made of two other approaches. First, there is a literature on uncovering network structure given covariances of outcomes. The state of the art in this work is Drton, Foygel, and Sullivant (2011) who examine global identification of the parameters  $a_{ij}$  for models of the form  $\omega_i = \sum_{j \neq i} a_{ij} \omega_j + \varepsilon_j$ . We discuss this work in Section

5.i. Second, Lee (2007) and Lee, Liu, and Lin (2009) use ideas from the spatial

Fortin provide a condition on the matrix defining social interactions which is sufficient for identification of the parameters in the model they consider. They also demonstrate how Manski's reflection problem can arise when their condition is not met. Blume, Brock, Durlauf, and Ioannides show that the identification is generic in a precise sense in this class of linear models. They also begin the exploration of the identification of social network effects when the weighting matrix is not known. This is certainly the case in most existing data sets. Finally, they provide an explicit microeconomic foundation for the linear-in-means model. They show that differences between the Manski and Bramouille, Djebbari, and Fortin results involve the use of approximations to the appropriate underlying Bayes-Nash equilibrium that produces linear behavioral equations. Both of these papers argue that the social interactions models that have been the basis of the existing econometric literature are a special case of a general social networks structure.

Although research on the identification problem has begun, a systematic investigation of the issues facing an empiricist has yet to be undertaken. Four examples serve to illustrate this gap. First, little thought has been given to distinct transmission mechanisms for endogenous social influence (i.e. the influence of expected behaviors of others on a given individual's actions), and for contextual social influence (the influence of exogenous characteristics of others on a given individual's choices). Surprisingly, existing models assume these mechanisms are sufficiently similar that they can be described by the same matrix of social weights. There is no reason why this should be so and it is easy to imagine cases where the networks would differ. Within a classroom, conformity effects may lead students to be more sensitive to the effort of those students of like ethnicity and gender, while the desire to perform well relative to the class distribution may induce a different contextual effect on effort based on past performance of other students. Models that make distinctions between the different

---

statistics literature to model social networks and address identification problems pertaining to social factors. This work takes a much more restrictive view of networks than we do because of the assumption that agents are located in an associated spatial interaction structure that generalizes the notion of a Markov process. Conley and Topa (2002) propose ways of measuring proximity in social space, but construct measures that are used to test for spatial dependence, rather than measure social influences in the way in which we conceptualize them.

endogenous and contextual-effects transmission mechanisms are better grounded in theories of social influence and can have different statistical properties that provide alternative paths to identification.

Second, the impact of endogenous network construction is rarely addressed in systematic fashion. Structural models of network formation are typically not linked to behavior within networks. Instead network endogeneity is addressed using instrumental variables are employed whose validity is often unclear when one considers network formation and behavior in networks as two stages of a game.

Third, there is, to the best of our knowledge, there exists no systematic discussion of the informational content of aggregated data for network effects. The few studies examining aggregate data, most notably Glaeser, Sacerdote and Scheinkman (1996) and Graham (2008), focus on the use of aggregate data to estimate particular parameters or provide evidence of some type of social effect rather than assess overall information content.

Finally, while Blume, Brock, Durlauf and Ioannides provide examples of identification under partial network observability, this question has received very little attention in econometrics. A partial exception is Conley and Topa (2003) who explore mismeasurement of groups in social interactions models.

This paper provides a systematic analysis of identification in linear social networks models. This is both a theoretical and an econometric exercise in that it links identification analysis to a rigorously delineated model of interdependent decisions. The paper proceeds as follows. Section 2 describes a social network game from which the linear model emerges as a unique equilibrium. This section introduces different mechanisms for the spread of endogenous and contextual effects through the network. Section 3 provides the conceptual framework which we use to study identification. Section 4 explores identification results when network structure is both exogenous and known *a priori*. It considers how differences in the spread of contextual and peer effects can aid identification. It also addresses identification from aggregate data. Section 5 considers identification when the network is exogenously given but is not observed by the econometrician. Section 6 examines endogenous network formation. The theoretical model of Section 2 is extended to a simple network formation game. Two

different models of preferences for networks are discussed, and identification conditions are developed. The possibility of extending control function techniques to account for network endogeneity is discussed. Section 7 concludes. A technical appendix follows which contains all proofs.

## **2. A social networks game**

In this section we provide an explicit derivation of a linear social networks model for individual behavior as the *Bayes-Nash equilibrium* for a social networks game and demonstrate that these linear behavioral rules are the unique descriptions for individual behavior. An analogous result is developed for social interactions models in Blume, Brock, Durlauf, and Ioannides (2010). As is the case for the earlier derivation, the unsurprising key to justifying linear social networks models as econometric specifications for individual behavior is to assume that individual agents possess quadratic payoff functions. As such our model is a species of quadratic interaction games that have become popular in different contexts.

### **i. the quadratic social networks model**

The social networks game we describe is a game of incomplete information. In this game, each individual is described by a bundle of characteristics, some observed by everyone, including the econometrician, some observed by the individuals in the population but not by the econometrician, and some private to the individual. Individuals have preferences over their actions, which are the sum of a private component and a social component.<sup>5</sup> The private component, which is quadratic in an individual's own actions, varies across individuals. Some part of the variation is common knowledge, and some is private. The social component is common to all individuals. Each

---

<sup>5</sup>Game-theoretic models are usually interpreted to have individual preferences over outcomes, which are jointly determined by player actions. Alternatively, they may be viewed as models of externalities, where individuals' preferences over their own choices are mediated by the decisions of other players. Here we adopt the latter view.

individual's utility is decreasing in the distance between his action and a weighted average of the actions of those who influence him. The equilibrium concept is Bayes-Nash: individuals choose an action to maximize their expected utility given their information about themselves and the public information about everyone in the population. Equilibrium beliefs are constructed from the individual's strategy functions and the common prior belief. Our assumptions will imply that equilibrium strategies are linear decision rules of the type that are standard in the empirical literature.

The population of network participants, the set of players, is a set  $V$  containing  $|V| < \infty$  members. Each individual is described by a vector of characteristics in  $\mathbf{R}^{P+2}$ , a vector  $(x_i, v_i, \varepsilon_i)$  where  $x_i \in \mathbf{R}^P$  is a vector that is observable to all network participants and to the (presumably external) econometrician,  $v_i \in \mathbf{R}$  is observable to the network but not to the econometrician, and  $\varepsilon_i \in \mathbf{R}$ ,  $i$ 's private type, is observable only by himself. In Bayesian games, individuals are described by *types*, which detail who they are and what they know. The vector  $(x_i, v_i, \varepsilon_i)_{i \in V} \in \mathbf{R}^{|V|(P+2)}$  is the *state* of the game. The *type* of individual  $i$  when the game is in this state is the vector  $t_i = (x, v, \varepsilon_i)$ .<sup>6</sup> Individual  $i$ 's type reflects his characteristics and the public knowledge he observes, namely, the  $x_j$ 's and  $v_j$ 's of other individuals. The *a priori* distribution of game states is exogenous, and is described by a probability distribution  $\rho$ . Knowledge of  $\rho$  is common to all individuals.

Each individual chooses an  $\omega_i \in \mathbf{R}$ . Individual  $i$ 's utility is a function of his type, his action, and the actions of others in the population. His payoff function is

$$u_i(\omega_i, \omega_{-i}) = \left( \sum_{p=1}^P \left( \gamma^p x_i^p + \delta^p \sum_j c_{ij} x_j^p \right) + v_i + \varepsilon_i \right) \omega_i - \frac{1}{2} \omega_i^2 - \frac{\phi}{2} \left( \omega_i - \sum_{j \neq i} a_{ij} \omega_j \right)^2. \quad (1)$$

---

<sup>6</sup>A note on notation: for any individual-level vector  $z_i$ , the unsubscripted variable  $z$  refers to the vector  $(z_i)_{i \in V}$  and  $z_{-i}$  refers to the vector obtained from  $z$  by omitting  $z_i$ .



This payoff function takes as special cases many of the payoff functions that have at least implicitly appeared in the literature. As far as we know our analysis is the first fully rigorous demonstration of existence and uniqueness of an equilibrium for a general quadratic social networks model.

The first two terms constitute the component of the payoff function that is independent of the choices of others; we call this the private component of the payoff function. The marginal value of the individual's choice depends upon his characteristics and a weighted average of the characteristics of others, computed with the weights  $c_{ij}$ . The matrix of these weights is  $C$ . This matrix is assumed to be nonnegative and each row sums to 1 this generalizes the common assumption in the social interactions literature that the unweighted average of individual characteristics affects each member of a group. The final component captures a purely social component to payoffs in that the component depends on the choices of others. It is quadratic and decreasing in the squared distance between the individual's choice and a weighted average of the choices of others, computed with weights  $a_{ij}$ . The matrix of these weights is  $A$ . The parameter  $\phi$  determines the weight placed on the public component relative to the private component. Accordingly,  $\phi \geq 0$ . Note that if the values of the  $c_{ij}$ 's and the  $a_{ij}$ 's are known *a priori*, there are  $2P + 1$  utility parameters which determine choice.

The private component includes a conventional externality, that network average characteristics are a kind of group capital or public good. This is expressed in the term

$$\sum_{p=1}^P \delta^p \sum_j c_{ij} x_j^p.$$

The literature refers to this term as the *contextual effect*. Here we have extended the idea from groups, and group averages, to networks. Contextual effects for the entire network are thus summarized by the sociomatrix  $C$ .<sup>7</sup> The literature on social interactions presents two kinds of contextual variables: group averages of individual level variables, and distinct group variables. In a study of peer effects on educational outcome, for instance, classroom-specific average family income and teacher-specific variables are examples of the first and second type, respectively. It has been known

---

<sup>7</sup>Sociomatrix is a term from sociology. In the mathematics literature these objects are called weighted adjacency matrices.

since Brock and Durlauf (2001a,b) that the relationships between these variables determine whether identification holds. In particular, they show in a related model that the presence of one individual variable whose group average is not a contextual variable is sufficient for identification. Since they raise no new issues here, we omit group variables which are not averages of individual variables.

The social component captures the idea that deviating from the average behavior of one's peers is costly. This is the source of so-called peer, or endogenous, effects. As in the case of the contextual effects, the weights  $a_{ij}$  extend uniform group averages, which underlie social interactions models to more general social networks, and different weighting schemes. Hence, we assume  $a_{ij} \geq 0$  and that the rows of  $A$  sum to 1. Whereas the empirical literature uses simple group averages as a model of peer effects, the motivation for peer effects is that they measure social influence. Thus the peer-effects network should not have self-loops. Accordingly, we assume that the  $a_{ii}$ 's are 0. This difference is pertinent for the identification of utility parameters in the linear-in-means model.

In the existing econometric literature, the same sociomatrix is used to average endogenous and exogenous variables. Here we consider other schemes. To see why this is plausible, consider again peer effects on educational outcomes. Suppose that peer effects really are from peers. In this case, the sociomatrix averaging endogenous effects should measure friendships or social influence. On the other hand, variables such as average family income may work at the classroom or school level. Models like this have two social networks: the peer effects network through which endogenous interactions are transmitted, and the contextual effects network, which determines the contextual effects, each represented by its sociomatrix,  $A$  and  $C$ , respectively. Notice that individuals do not need to know the entire networks. They need know only to whom they are connected and the weights assigned to them. In this sense, each individual has a payoff-relevant neighborhood.

Since we are working in an environment with private types, we need to make some assumptions on unobserved and observed individual-specific heterogeneity. In the following definition,  $x$  is an  $|V|P$ -dimensional vector in which, for each  $i$ , the

observations  $x_i$  are stacked variable by variable, with the individual-specific vectors so created stacked as well.

Our various assumptions may be summed up by the following axioms.

**Axiom 1.**  $\phi \geq 0$ ,  $A$  and  $C$  are non-negative, their row sums are 1, and for all  $i$ ,  $a_{ii} = 0$ .

**Axiom 2.** Second moments of the marginal distribution  $\rho^\varepsilon$  exist.

**Axiom 3.** Second moments of the marginal distribution  $\rho^v$  exist.

**Axiom 4.** For all  $i$ ,  $E(\varepsilon_i | x, v, \varepsilon_{-i}) = \mu_i^\varepsilon$  is independent of  $x$  and  $\varepsilon_{-i}$ .

Axiom 1 restricts the payoff function.<sup>8</sup> The second and third axioms guarantee that the choice problems required of individuals by the game are well-posed, that the necessary expectations exist. The joint distribution  $\rho$  on  $x$ ,  $v$  and  $\varepsilon$  is the common prior belief on the space of types. If Axiom 4 were false, the equilibrium strategies need not be linear in  $x$ .

A strategy for individual  $i$  is a function that assigns an action to each of his possible types, a function  $f_i : \mathbf{R}^{V(P+1)+1} \rightarrow \mathbf{R}$ . A Bayes-Nash equilibrium BNE of the game is a strategy profile  $(f_i)_{i \in V}$  such that each  $f_i$  maximizes  $E(u_i(\omega_i, \omega_{-i}))$  where the expectation is taken with respect to the strategies  $f_{-i}$  and the common prior  $\rho$ .

**Theorem 1.** Assume the sociomatrices satisfy Axiom 1. For any prior distribution  $\rho$  satisfying Axioms 2 and 3, there exists a unique BNE. The equilibrium strategy profile can be written in the form

---

<sup>8</sup>For an example that does not fit into our framework, see Calvó-Armengol, Patacchini and Zenou (2009).

$$\left( f_i^*(\mathbf{x}, \nu, \varepsilon_i) \right)_{i=1}^{|M|} = \frac{1}{1+\phi} \left( I - \frac{\phi}{1+\phi} \mathbf{A} \right)^{-1} \left( \nu + \mu^\varepsilon + \sum_{p=1}^P (\gamma^p I + \delta^p \mathbf{C}) \mathbf{x}^p \right) + \frac{1}{1+\phi} \varepsilon_i^{dev} \quad (2)$$

where  $\varepsilon_i^{dev} = \varepsilon_i - \mu_i^\varepsilon$ .

Notice that strategy profile implies distinct roles for  $\mu_i^\varepsilon$  and  $\varepsilon_i^{dev}$  since only the former is common knowledge.

The equilibrium strategies map types into actions; that is, strategies describe a map  $f : (\mathbf{x}, \nu, \varepsilon) \rightarrow \omega$ . This is what the empirical literature calls a reduced form. This is, in fact, the structural model for a game theorist insofar as what one means by a structural model is a model delivered from theoretical considerations. In the social networks and social interactions literature, what are called structural models are equations in which individual choices are determined by the individual choices (or expected choices) and characteristics of others as well as the characteristics of individuals. At best, these may be first-order conditions. In our model, for example, the first order conditions for utility maximization are, for all  $i$ ,

$$\omega_i = \sum_{p=1}^P \frac{\gamma^p}{1+\phi} \mathbf{x}_i^p + \sum_{p=1}^P \frac{\delta^p}{1+\phi} \sum_j \mathbf{c}_{ij} \mathbf{x}_j^p + \frac{\phi}{1+\phi} E \left( \sum_j \mathbf{a}_{ij} \omega_j \middle| t_i \right) + \frac{\nu_i + \varepsilon_i}{1+\phi} \quad (3)$$

which, are necessary conditions for maximization satisfied by the equilibrium strategies. Since the empiricist typically transforms these first order conditions in order to eliminate direct dependence on others' choices (or beliefs), he refers to equation (3) as structural model for  $i$ , and equation (2) as the reduced form for the system.

Regardless of the different perspectives, an interesting econometric exercise is to determine the parameters that describe the utility function and the network, for a variety of purposes, including the exploration of positive and normative effects of policies. To avoid confusion, we will refer to the right hand side of equation (2) neither

as the structural model (which it is) or a reduced form (which it is often called, but is not), but instead as the strategy profile of the network. We will abuse this term slightly, because we will also use the term strategy profile to refer to the matrix  $B = (B_1, \dots, B_p)$  in which each matrix  $B_p$  acts on the characteristics  $x^p$  and is given by the function  $B_p(\gamma, \delta, \phi, C, A)$  defined as follows:

$$B_p = B_p(\gamma, \delta, \phi, C, A) = \frac{1}{1 + \phi} \left( I - \frac{\phi}{1 + \phi} A \right)^{-1} (\gamma^p I + \delta^p C). \quad (4)$$

Notice that Axiom 1 implies that the matrix inverse on the right hand side exists for all  $\phi \geq 0$ . From equation (2) it can be seen that the matrix  $B$  essentially characterizes the equilibrium strategies, and our identification exercises essentially involve determining what parameters can be recovered from them. If the span of  $x^p$  has dimension less than  $|V|$ , then  $B_p$  will be unique only up to its action on a lower-dimensional subspace. But this is an issue of identification, not one of existence or uniqueness.

## ii. **social interactions models as special cases of the general linear social networks model**

The social interactions literature has focused on the special case where individuals react to the average of others in a predefined group  $g$ . Notationally,  $g$  denotes a collection of indices corresponding to population members. Social interactions models assume that each member of a group is affected by the average behavior of others in the group and is unaffected by individuals who are not members of the group. Following Blume, Brock, Durlauf, and Ioannides (2010), the microfounded quadratic social interactions model is a special case of our social networks model such that<sup>9</sup>

---

<sup>9</sup>Bramoullé, Djebbari, and Fortin refer to the case where  $i$  is omitted from the averaging as exclusive averaging and associate this form of averaging with Moffitt (2001). They

$$\begin{aligned}
c_{ij} &= \frac{1}{n_g} && \text{if } i, j \in g, \\
a_{ij} &= \frac{1}{n_g - 1} && \text{if } i, j \in g, i \neq j \\
c_{ij} = a_{ij} &= 0 && \text{otherwise}
\end{aligned} \tag{5}$$

where  $n_g$  is the size of group  $g$ . Under these restrictions, the first order condition for an individual's choice produces the first order condition (interpreted as a structural equation in the literature)<sup>10</sup>

$$\omega_i = \sum_{p=1}^P \frac{\gamma^p}{1+\phi} x_i^p + \sum_{p=1}^P \frac{\delta^p}{(1+\phi)n_g} \sum_j x_j^p + \sum_{j \neq i} \frac{\phi}{(1+\phi)(n_g-1)} E(\omega_j | \mathbf{x}) + \frac{1}{1+\phi} \varepsilon_i \tag{6}$$

When the population size is large, this expression becomes arbitrarily close to

$$\omega_i = \sum_{p=1}^P \frac{\gamma^p}{1+\phi} x_i^p + \sum_{p=1}^P \frac{\delta^p}{(1+\phi)} \bar{x}_g^p + \frac{\phi}{(1+\phi)} E(\bar{\omega}_g | \mathbf{x}) + \frac{1}{1+\phi} \varepsilon_i \tag{7}$$

---

contrast this with inclusive averaging, in which  $i$ 's behavior is included when averaging, associating this form of averaging with Manski (1993). In our view, inclusive averaging does not make behavioral sense for endogenous effects. We believe that the correct interpretation of Manski's approach, confirmed in conversation with him, is that his formulation was based on the assumption that the group size was large enough that own effects on averages could be ignored. In contrast, inclusive averaging can make behavioral sense when contextual effects derive from public goods, for example. Hence, (5) is the appropriately microfounded social interactions analog to the social networks model. We are unaware of any work in the social interactions literature that has allowed for exclusive and inclusive averaging to coexist in the same population.

<sup>10</sup>We omit  $\nu$  from this specification since this term does not appear in social interactions models.

where the barred variables are group averages. This last equation defines the linear-in-means model that has received so much attention in the econometric literature.

### 3. Identification concepts

In this section we provide the assumptions we append to the theoretical model as we move from theoretical to econometric issues. We elucidate what we mean by identification, how identification results depend on an analyst's objective, and the relationship between identification notions and data moments.

#### i. basic ideas

Identification is concerned with the problem of making relevant distinctions between different parameter values based on some observables and some *a priori* knowledge of the data-generating process. In order to do this, one must specify the following objects: The set of structures that could conceivably have generated the data; a statistic (for example, a sample moment), from which the econometrician will infer structure; *a priori* knowledge of the econometrician, which imposes restrictions on the set of conceivable structures which are made *ex ante* the observation; and finally, a description of the distinctions among the structures that the econometrician would like to draw. A *structure*  $m$  is a description of a data generating process in terms of parameters, some of which are of interest to the econometrician. A *model*  $M$  is the set of conceivable structures. Each structure  $m \in M$  generates a probability distribution  $\Delta(m)$  on the set of values of the statistic. *A priori* information is represented, as knowledge usually is, by an information partition — in this case, of the set  $M$  of models.

The specification of a structure may be quite complicated, and the econometrician might be interested only in some part of the structure; parameters, for instance, that may be tuned by policy changes. The idea that we need to identify, and only identify, useful knowledge is as old as the notion of structural econometric models. Heckman (2000,2005) has reminded us of the importance of this idea, which he calls

*Marschak's Maxim* in acknowledgement of its earliest clear statement in Marschak (1953). Useful knowledge can also be represented by a partition, or equivalence relation, on the model  $M$ . Two structures are equivalent in the sense of Marschak if they differ only in ways that are not of interest to the econometrician. For example, in the model of Section 2 with exogenously given networks, no conceivable statement about the positive or normative implications of some policy experiment will involve anything more than means and variances of the common prior  $\rho$ , so there is no point in trying to identify, say, the third moments of  $\rho$ . One way in which Marschak equivalence arises is when one asks if particular parameters are identified. To investigate the identification of parameter  $p$  is (at least implicitly) to regard as equivalent structures which share a common value of parameter  $p$ , regardless of how else they may differ.

Identification has to do with the inverse image of  $\Lambda$ . Formally, if Marschak equivalence is represented by an equivalence relation  $\sim$  on the set  $M$  of models, and if  $S$  is the set of structures that are possible given the econometrician's ex ante information, then identification is achieved if for any two different distributions  $q'$  and  $q''$  of the statistic,  $s'$  and  $s''$  are structures in  $\Lambda^{-1}(q') \cap \Lambda^{-1}(q'') \cap S$ , then  $s' \sim s''$ . Claims about identification are often confused with claims about inference; for example, that if a parameter is identified, then it can be consistently estimated. On the contrary, identification is concerned with claims about the population distribution of data rather than with the existence of statistics with certain inferential properties.

We now apply these ideas to the linear social network model of Section 2

**Definition 1.** A structure is a list  $\langle (\gamma^p, \delta^p)_{p=1}^P, \phi, C, A, \mu^\varepsilon, \mu^v, \rho \rangle$ . A model is a set of structures satisfying Axioms 1 through 4. Denote by  $M$  the set of structures with the following properties:

- i. The span of  $x$  has dimension  $|V|P$ .
- ii. For all  $i$ ,  $E(v_i | x) = \mu_0^v$ , is independent of  $x$  and  $i$ .



- iii.  $\mu_i^\varepsilon$  is independent of  $i$ .
- iv. For all  $i$  and  $j$ ,  $a_{ij} > 0$  if and only if  $a_{ji} > 0$ .
- v. For all  $i$  and  $j$ , there is a pair  $i \neq j$  with  $c_{ij} > 0$ , and  $c_{ij} > 0$  if and only if  $c_{ji} > 0$ .
- vi. One of  $\delta$  and  $\gamma$  is not 0.

These properties are further assumptions on the theoretical model that address identification issues rather than existence and uniqueness of the equilibrium. Recall that we have already assumed that  $\mu^\varepsilon$  is independent of  $x$ , and that the  $a_{ij}$  are 0. Condition i ensures that  $B$  is unique, that the relevant space on which strategies are defined is full-dimensional. Conditions ii and iii shrink the size of the parameter space considerably, and can be justified as consequences of assumptions such as exchangeability that may be employed in any event. Conditions iv and v of the definition impose the restriction that the location of 0's in the sociomatrices is symmetric in a weak sense. That is,  $i$  influences  $j$  if and only if  $j$  influences  $i$ . This is done for technical convenience. Notice that the influence weights can be quite different, so disallowing one-way influence rules out only boundary cases. We rule out the identity matrix. The purpose of condition vi is to rule out a degenerate case: if  $\delta = \gamma = 0$ , then  $\omega_i$  is determined only by  $\varepsilon$ , and in this case  $\phi$  cannot be identified without further assumptions on the (joint) distribution of the private types.

*A priori* information in this paper will mostly have to do with parameter values. Section 4 is concerned with identification when both sociomatrices  $A$  and  $C$  are known *a priori*. Section 5 investigates the degree to which this assumption can be relaxed. In neither case is *a priori* knowledge of the common prior  $\rho$  necessary for identification. When network formation is endogenous this is no longer the case. At different points we

will assume that first moments of  $\rho$ , conditional moments of  $\rho$ , and  $\rho$  itself are all *a priori* knowledge. One use of *a priori* knowledge assumptions is to define particular parameter restrictions. For instance, in some of our theorems it is known *a priori* that  $\delta = 0$  (no contextual effects). We will also be interested in generic identification. This too can be expressed in terms of *a priori* knowledge.

**Definition 2.** The parameter  $\rho$  in model  $M$  is generically identified from the joint distribution of  $\omega$  and  $x$  if and only if there is a closed and lower-dimensional set  $M^{crit}$  such that if the complement of  $M^{crit}$  is known *a priori*, then  $\rho$  is identified.

The proofs identify how this set can be computed in any given instance, but we will not report on the description of the so-called “critical set” of models where identification may fail.

Useful knowledge in this paper will be concerned with identifying which parameters or functions of parameters are identified. For instance, it is usually the case that  $\gamma + \delta$  is identified. By this we mean that the model space can be partitioned so that two models are in the same element of the partition (the two models are equivalent) if and only if the sum  $\gamma + \delta$  is the same for both models.

In the next section and in the remainder of the paper, we will incur no loss of generality and gain greater clarity by taking  $P = 1$ , that is, from assuming there is only one exogenous variable.

## ii. sample moments and identification

Given the axioms in Section 2 and the requirements imposed on a model, the conditional distribution of  $\omega$  given  $x$  is described by an equation of the form

$$\omega = B(m)x + \mu^e + \mu^v + v^{dev} + \varepsilon^{dev} \quad (8)$$

In this way the parameters  $\mu^\varepsilon + \mu^\nu$ ,  $\gamma$ ,  $\delta$ ,  $\phi$ ,  $A$  and  $C$ , and the distribution of  $\nu^{dev}$  and  $\varepsilon^{dev}$ , the deviations of  $\nu$  and  $\varepsilon$  from their means, completely determine the conditional distribution of  $\omega$  given  $x$ ; this and the marginal distribution of  $x$  determines the distribution of the pair  $(x, \omega)$ . The identification question is whether one can recover these parameters from a given joint distribution of  $\omega$  and  $x$ .

One can immediately make a couple of observations. First, given a joint distribution, the matrix  $B$  can be recovered. Then the difference  $E(\omega|x) - Bx$  identifies the sum  $\mu^\varepsilon + \mu^\nu$ , and this is the best that can be done-these parameters cannot be separated. Another observation comes from equation (4). Since the row sums of  $A$  are all 1, so are the row sums of  $(1+\phi)^{-1}(I-\phi(1+\phi)^{-1}A)^{-1}$ . Since the row sums of  $C$  are 1, it follows that for any  $B \in B(m)$ , the row sums of  $B$  are  $\gamma + \delta$ . The following is immediate.

**Lemma 1.** The sum  $\gamma + \delta$  and the sum  $\mu^\varepsilon + \mu^\nu$  are identified in  $M$  from the joint distribution of  $\omega$  and  $x$  without any additional *a priori* information.

While  $B$  is always observable from individual data, it may not be observable from aggregate data. We discuss this in Section 3.iv.

Most often, equation (8) is estimated with a regression model; that is,  $E(\omega|x)$  is the object of statistical enquiry, and identification strategies involve the recovery of the parameters from this conditional mean. However, there are other approaches. The fact that complementary network connections create correlation between actors resulting in excess variation is an old observation in network science (e.g. Ising (1925) and Dobrushin (1965)). It was first exploited in econometric models by Glaeser, Sacerdote and Scheinkman (1996,2003) and subsequently by Graham (2005,2008) and others. It is occasionally alleged that investigation of  $E(\omega^2|x)$  creates new opportunities for identification. This claim is true but limited. It is apparent from eq. (8) that getting

something from variances requires characteristics known to the public but not to the econometrician

A particular instance of this is Graham's (2005,2008) variance-contrast method. Graham (2005) provides a theoretical foundation for this econometric model, which is a game of complete information but with individual characteristics unobservable to the econometrician. Imposing the parameter restrictions  $\delta = \gamma = 0$  in the game of section 2 gives a game of complete information where characteristic  $\nu$  is observed by individuals but not by the econometrician. (For Graham realized choices are observed by each agent.) Assuming further  $A = C$  and that the social network is the union of cliques of different sizes and that each individual weighs everyone in his group equally, Graham establishes identification of  $\phi$ .

Aside from this type of information structure, however, variance and higher-moment methods add nothing to the possibilities for identification. Variances augment identification if there are some variables that  $B(m)$  acts upon and which the econometrician does not observe. Otherwise the econometrician is left with  $\text{var}(\varepsilon)$  which reveals little of interest.

#### **4. Identification with known sociomatrices**

In this section we consider identification when both  $A$  and  $C$  are both known *a priori*. The goal here is to study the traditional reflection problem, that is, identification in the presence of contextual effects. If  $A = C$ , the analysis with individual-level data is a straightforward extension of Blume, Brock, Durlauf, and Ioannides (2010). The discussion of aggregation is new.

##### **i. identification with individual-level data**

Recall that the sums of the means of the two unobserved variables are identified in this and all subsequent models. This is a trivial point in light of the fact that the sum of

the means of the unobservables is nothing more than the constant term in the individuals' strategies.

The following lemma is useful for checking identification when the peer- and contextual-effects network stand in particular relationships to each other.

**Lemma 2.** If  $A$  and  $C$  are *a priori* knowledge, if  $B(m') = B(m'') = B$  and if there is a pair  $i \neq j$  such that  $c_{ij} = 0$  and  $b_{ij} \neq 0$ , then  $(\gamma', \phi', \delta') = (\gamma'', \phi'', \delta'')$

The next theorem gives an example of how the lemma can be employed, giving one set of assumptions on the primitives which fulfill the conditions of the lemma.

**Theorem 2.** Suppose that the following facts are known *a priori*:

- i.  $A$  and  $C$ ;
- ii. the peer-effects network is connected;
- iii. there is no individual  $j$  such that  $c_{kj} = 0$  for all  $k$
- iv. there is some pair  $i, j$  such that  $c_{ij} = 0$ .

Then  $\gamma$ ,  $\delta$ , and  $\phi$  are generically identified from the joint distribution of  $\omega$  and  $x$ .

There are many ways of extending this theorem to multiple peer effects cliques, especially if the peer- and contextual-effects cliques are not nested.<sup>11</sup> In particular, individuals who influence a peer-effects clique through a contextual effect but are not themselves part of the clique identify peer effects in a manner analogous to the Brock and Durlauf (2001b) condition for identification, which requires the existence of an

---

<sup>11</sup>A clique is a network in which all individuals are connected, i.e. all off-diagonal elements of the sociomatrix are positive.

individual variable whose group average is not a contextual variable. This requires some structure on the contextual effects network. The network is already assumed to be bidirectional, that  $i$  has a contextual effect on  $j$  if and only if  $j$  has a contextual effect on  $i$ . We will assume in addition that the contextual effects network is transitive: If  $i$  is affected by  $j$ , and  $j$  is affected by  $k$ , then  $i$  is directly affected by  $k$ . Formally, if  $c_{ji} > 0$  and  $c_{jk} > 0$ , then  $c_{ik} > 0$ .

**Theorem 3.** Suppose that the following facts are known *a priori*:

- i.  $A$  and  $C$ ;
- ii. the contextual effects network is transitive;
- iii. there are components  $V_1^C$  and  $V_2^C$  of the contextual-effects network and component  $V^A$  of the peer-effects network such that each  $V_i^C \cap V^A \neq \emptyset$ .

Then  $\gamma$ ,  $\delta$  and  $\phi$  are identified from the joint distribution of  $\omega$  and  $x$ .

Bramoullé, Djebbari, and Fortin (2009) provide a powerful identification requirement for the traditional linear-in-means model that provides a connection between identification and network structure. The next result extends this to our two-sociomatrix model.

**Lemma 3.** Suppose that  $A$  and  $C$  are known *a priori*.

- i. Suppose it is known *a priori* that  $A \neq C$  and  $AC \neq A, C$ . The matrices  $I$ ,  $A$ ,  $C$  and  $AC$  are linearly independent if and only  $\gamma$ ,  $\delta$  and  $\phi$  are identified from the joint distribution of  $\omega$  and  $x$ .

- ii. Suppose it is known *a priori* that  $A \neq C$  and  $AC = C$  and that the matrices  $I$ ,  $A$ , and  $AC$  are linearly independent. Then  $\gamma$ ,  $\delta$  and  $\phi$  are identified from the joint distribution of  $\omega$  and  $x$ .
- iii. Suppose it is known *a priori* that  $A = C$  and that  $\gamma + \delta \neq 0$ . Then *a priori* knowledge that  $I$ ,  $A$ , and  $A^2$  are independent is necessary and sufficient for  $\gamma$ ,  $\delta$  and  $\phi$  to be identified from the joint distribution of  $\omega$  and  $x$ .

The condition that  $\gamma + \delta = 0$  ensures that peer and contextual effects do not cancel each other out.

It can be shown that the set of pairs of sociomatrices failing to satisfy the independence condition Lemma 3.i is closed and lower-dimensional in the space of all sociomatrices satisfying our requirements. Dependence is the existence of a non-zero solution in  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  of the following equation system:

$$\begin{aligned} \alpha + \beta c_{ii} + \delta \sum_l a_{il} c_{li} &= 0 \text{ for all } i, \\ \beta c_{ij} + \gamma a_{ij} + \delta \sum_l a_{il} c_{li} &= 0 \text{ for all } i \neq j. \end{aligned} \tag{9}$$

Various cases of this system can be used to generate any number of conditions guaranteeing identification of  $M$  when  $A$  and  $C$  are known *a priori*. Here is one such instance, in which the separation of peer and contextual effects exposes yet another way in which the reflection problem is fragile.

**Corollary 1.** Suppose  $A$  and  $C$  are known *a priori*, and also that the contextual effects network is a clique, and in it all weights are equal, and that there exist two pairs of individuals  $i \neq j$  and  $k \neq l$  such that  $a_{ij} \neq a_{kl}$ . Then  $\gamma$ ,  $\delta$  and  $\phi$  are identified from the joint distribution of  $\omega$  and  $x$ .

## ii. aggregation

Classroom-level and village-level data often come aggregated. For example, an education data set may contain observation on mean outcome and mean characteristics of many classrooms. What can be learned in this case? The answer is, not very much. In general, with only mean characteristics and mean outcome data, identification will be complicated by the fact that there is no direct relationship between mean characteristics and mean outcome. A special case where there is such a relationship arises when the sociomatrices are bistochastic, that is column sums as well as row sums all equal one. Although this is a very restrictive condition, it includes the important case of equal-weighted averages of all other individuals. This is the case which Graham (2008) uses to show how  $\phi$  may be recovered from the variance in mean group outcomes, if one can calculate this variance for different sized groups.

Suppose there are  $N$  observational units such as classrooms or villages and unit  $n$  has member set  $V^n$ . We suppose that if individual  $i$  is in unit  $m$  and individual  $j$  is in unit  $n \neq m$ , then  $a_{ij} = c_{ij} = 0$ ; the units are not connected in either the peer- or contextual-effects social network. Observational units are identified with superscripts. The  $n$ 'th observational unit has peer- and contextual-effect sociomatrices  $A^n$  and  $C^n$  respectively. (Note that  $N$  may equal 1.) Let  $e^n$  denote the vector where each element is  $1/|V^n|$ , where  $|V^n|$  is the number of individuals in observational unit  $n$ . The econometrician observes the averages  $(e^n \cdot x^n, e^n \cdot \omega^n)_{n=1}^N$ .

### a. bistochastic sociomatrices

As observed above, a non-negative matrix is bistochastic if its row sums and columns sums are both 1. One such matrix is the sociomatrix wherein each individual equally weights all other individuals. More generally, if the network is regular,<sup>12</sup> the weighted adjacency matrix that assigns equal weights to all individuals whose weights

---

<sup>12</sup>All nodes have the same degree.



are non-zero will be bistochastic. Inverses and products of bistochastic matrices are also bistochastic and in particular  $(1 + \phi)^{-1} (I - (\phi/(1 + \phi))A^n)^{-1}$  is bistochastic. As before, let bars represent group averages. Let bars represent group averages. The average outcome in group  $n$ , is

$$\begin{aligned} \bar{\omega}^n = & \\ e^n \mu^\varepsilon + e^n \cdot \frac{1}{1 + \phi} \left( I - \frac{\phi}{1 + \phi} A^n \right)^{-1} (\gamma x^n + \delta C^n x^n + \nu) + \frac{1}{1 + \phi} e^n \cdot \varepsilon^{dev} = & \\ \mu_0^\varepsilon + \bar{\nu} + (\gamma + \delta) \bar{x}^n + \frac{1}{1 + \phi} \bar{\varepsilon}^{dev}. & \end{aligned}$$

The obvious result is that only the sums  $\mu_0^\varepsilon + \mu_0^\nu$  and  $\gamma + \delta$  are identified.

**Theorem 4.** If only  $A$  and  $C$  are known *a priori*, then only  $\gamma + \delta$  and  $\mu_0^\varepsilon + \mu_0^\nu$  are identified from the joint distribution of groups, average choices, and average characteristics. No other parameters are identified.

We saw in Section 3.ii that these parameter sums are identified with individual observations, but this result is not generally true for models when only aggregate data are observed.

## b. exchangeable individuals

The analysis of identification from aggregate data for more general classes of social networks will require assumptions on the relationship between characteristic means and the distribution of characteristics among the population. Suppose that the distribution of characteristics has the following property:

**Property P:**  $E(x^n | e^n \cdot x^n = z) = (z, \dots, z)$ .

Property P says that the conditional mean characteristic of each group member given the group sample mean equals the group sample mean. This property follows if the  $x_i$ 's are exchangeable, among other hypotheses. The consequence of Property P is the following:  $E(\omega|x)$  can reveal some information, but it does not reveal  $B(m)$ . This may be seen in the calculation

$$E(\bar{\omega} | \bar{x}) = \left( \mu^\varepsilon + \mu^\nu + \frac{1}{1+\phi} \left( I - \frac{\phi}{1+\phi} A^n \right)^{-1} (\gamma I + \delta C^n) \right) \begin{pmatrix} E(x_1 | \bar{x}) \\ \vdots \\ E(x_{|V^n|} | \bar{x}) \end{pmatrix} + \begin{pmatrix} E(v_1^{dev} | \bar{x}) \\ \vdots \\ E(v_{|V^n|}^{dev} | \bar{x}) \end{pmatrix} = (\mu^\varepsilon + \mu^\nu) + (\gamma + \delta) \bar{x}$$

When Property P holds, the conditional expectation of the group mean choice given the group mean characteristic reveals nothing about parameter values. By implication, the approach to identification taken throughout this paper fails here. Notice, however, that whereas in the bistochastic case, the value of the random variable  $\bar{\omega}$  is independent of  $\gamma$ ,  $\delta$ , and  $\phi$ , here it is a conditional moment that fails to vary with parameters. This leaves open the possibility that other statistics may reveal the parameters. We have conducted some computations with variances and discovered that sometimes  $\phi$  can be identified from the conditional variance, and sometimes not. Clearly there is more work to be done here.

## 5. Identification with unknown social networks

### i. unknown peer-effects sociomatrices

In this section we consider cases wherein the contextual effects sociomatrix is a *priori* knowledge, but the peer-effects sociomatrix is unknown to the econometrician. Although it may seem surprising that one can identify the peer-effects sociomatrix conditional on knowledge of the contextual-effects sociomatrix, a moment's reflection shows why it is plausible. The dimension of the set of peer effects matrices is  $|V|(|V|-2)$ . The dimension of the set  $B((m:C=C^*))$  for a fixed  $C^*$  is no more than  $|V|(|V|-1)+1$ , but we can show it to be no less than  $|V|(|V|-2)$ . We need to recover  $|V|(|V|-2)+3$  parameters from  $B(m)$ . It is certainly plausible that a necessary order condition is satisfied. We are searching for sufficiency, however, and we will proceed by direct argument rather than by trying to pin down more carefully the structure of  $B(M)$ .

#### a. **identification without contextual effects**

It is common in the theoretical econometrics literature to assume that the peer-effects network is known, and in the empirical literature to pretend that it is. This is rarely the case, so it is important to see how far one can go without such knowledge. The first result concerns identification when it is known a priori that  $\delta = 0$ ; that there are no contextual effects. Our results differ from Drton, Foygel, and Sullivant (2011) because their analysis ignores individual and contextual influences on individual behavior which are at the heart of our analysis, because of our interest in generic as well as global identification, and because of the error structure we allow and the parameter constraints we impose as a consequence of the derivation of our behavioral equations from the Bayes-Nash equilibria we have described.<sup>13</sup>

Some empirical work in labor, public finance and health economics has been concerned with distinguishing peer and contextual effects. There are many natural economic problems, however, for which contextual effects create no identification

---

<sup>13</sup>Drton, Foygel, and Sullivant's interest in global rather than generic identification stems, among other reasons, from a concern about the properties of likelihood ratio statistics when a particular class of models is tested against a broader class. As far as we can tell, this is not an issue that naturally arises in economic contexts.

problem. Amazon's book pricing problem concerns networks. Are patterns of book demand due to homophily or a network effect? Price is a contextual variable, but since it is not an average of customers' characteristics, it creates no identification problems.

**Theorem 5.** If the econometrician knows *a priori* that  $\delta = 0$ , then the parameter  $\gamma$  is identified from the joint distribution of outcomes and characteristics. If  $\gamma \neq 0$ , then  $\phi$  and  $A$  are identified.

When  $\gamma = 0$ , all variation in outcomes is due to variation in the unobservable variables  $v_i$  and  $\varepsilon_i$ . With further assumptions, such as independence, made on these terms, we conjecture that the matrix  $A$  could perhaps be discerned.

This theorem claims that the utility parameter  $\phi$  governing the strength of the social interaction is identified when the peer-effects matrix is not known. Moreover, the peer-effects matrix itself is, in principle, recoverable from the data. This result is surprising to us, and the rest of this section will be concerned with how far this result can be pushed.

## b. identification with contextual effects

When contextual effects are present and the contextual-effects sociomatrix is known, it will turn out that parameter values are generically identified, and even when they are not,  $\phi$  is identified. We will assume it to be known *a priori* that  $\gamma I + \delta C$  is invertible. For a given  $C$ , the set of  $(\gamma, \delta)$  pairs for which invertibility fails is the union of a finite set of one dimensional spaces.

**Theorem 6.** If the sociomatrix  $C$  and the matrix  $\gamma I + \delta C$  both have full rank, are known *a priori*, then  $\gamma + \delta$  and  $\phi$  are identified from the joint distribution of outcomes and characteristics. There is a set  $\mathcal{C}$  of matrices whose complement in the set of all contextual-effects sociomatrices is closed and lower-dimensional, such that if  $C \in \mathcal{C}$ ,

then  $\gamma$ ,  $\delta$  and  $\phi$  are identified from the joint distribution of outcomes and characteristics.

It will be clear from the proof of this Theorem that if  $\delta$  and  $\gamma$  cannot be distinguished, the peer-effects network cannot be identified. Nonetheless, and this is the surprising feature of Theorem 6, the intensity of the peer group effect can still be measured.

It is worth noting that exclusion restrictions on  $A$  can create additional identification opportunities because it provides more equations with which to tie down  $\delta$ . In large social networks where each individual is connected to a small number of neighbors, a *priori* knowledge of the location of some  $A$ 's zeroes quickly leads to an over-identified system.

## ii. **identification with unknown peer- and contextual effects sociomatrices**

It should be clear that if both  $A$  and  $C$  are unknown to the econometrician, nothing is identified. In this case there are  $|V|(2|V|-3)+3$  parameters to identify, and the dimension of  $B(M)$  is at most  $|V|(|V|-1)$ . In this case, one faces the classic simultaneous equations identification problem (Fisher (1966), Hsiao (1983)). The conditions under which such systems are identified have long been well understood. Blume, Brock, Durlauf, and Ioannides (2010) give examples of linear and nonlinear coefficient restrictions that produce variants of identification. When the sociomatrices are sparse, as would occur in large networks wherein each individual has only a small number of connections, the resulting necessary and sufficient rank and order exclusion condition are likely to be easily met.

This last observation emphasizes the importance of survey data in identification of social network models when the analyst does not possess a *priori* knowledge of the network structure. It also indicates important limitations to current surveys. The AddHealth data set is arguably the most popular data set for the study of social network effects as it consists of a nationally representative sample of high school students who

are interviewed about their friends, among many other characteristics. Unfortunately, the data set's friendship questions are restricted in that each student is allowed to name up to 5 friends of each gender. Exclusion restrictions imply that it is more useful to know who is not someone's friend rather than who is. In other words, the AddHeath friendship questions, because they do not provide measures of friendship intensity, are best understood as distinguishing zero and nonzero elements in the sociomatrices for school populations. However, the restriction on the number of friends means that the failure to identify someone as a friend does not mean that there is a corresponding zero in the associated sociomatrices. While the limitation on the number of friends that could be named in the interviews has long been understood as inducing measurement error in network structure, as far as we know, the effects of this limitation on identification *per se* have not been recognized<sup>14</sup>.

One might hope that, as is the case with an unknown peer-effects, the magnitude of the peer effect might be identified. However, this is not the case

**Theorem 7.** Without *a priori* information,  $\gamma + \delta$  is identified from the joint distribution of actions and characteristics. The peer-effects parameter  $\phi$  is not identified.

## 6. Endogenous network formation

The endogenous creation of peer networks adds another layer of strategic complexity to the game of Section 2 and introduces a species of self-selection with all its attendant econometric issues. One of Heckman's seminal contributions to economics is the recognition that self-selection should not be treated as a nuisance, but rather as evidence that an additional behavior beyond the original one under study needs to be modeled. Here we explore the implications of endogenous network construction for the identification of utility parameters. In this section we will provide two

---

<sup>14</sup>Another concern is that the failure to identify someone as a friend is consistent with a negative entry in one or both of the sociomatrices we have employed. While we ruled this possibility out in our analysis, it obviously a possibility. We thank Jesse Naidoo for this observation.

results on parameter identification, differing in their hypotheses about *a priori* knowledge. We then discuss contemporary econometric techniques, control functions in particular but also other instrumental variable methods, within the context of these results. We describe how control functions may be useful for parameter estimation, and also point out some potential pitfalls in the choice of instruments for interactive decision problems like Bayesian games.

#### **i. a group membership game**

There is no one obvious network formation game to study, and so we will demonstrate the possibilities for selection in an extended example, a two-stage Bayesian game of group formation. Any strategic model of group formation must first ask, why do groups form? A distinguishing feature of social networks is the property of homophily, that similar individuals are attracted to one another. A large body of social science research (see McPherson et al. 2001) has documented that individuals in a social network are more likely to be directly connected to similar others. The urtext of sociological research on homophily, Lazarsfeld and Merton (1954), distinguishes two types of homophily, differing in their notions of similarity of individuals: *Status homophily* is the tendency for individuals to associate with those carrying similar markers of social status, such as age, ethnicity, gender, race, and income. *Value homophily* is the tendency of individuals associate with those who share common beliefs and values, regardless of their social status. The model we present below attempts to measure both of these pressures for affiliation. We capture this by modifying the payoff function of Section 2.

To make things concrete we will suppose that individuals can join one of two groups, *a* or *b*. (It may turn out that one group will be empty.) The strategic situation of Section 2 is extended to two stages. In the first stage, individuals observe all public information about types, and then simultaneously choose to join group *a* or group *b*, perhaps by walking to a particular location. At the second stage, individuals observe who is in their group, and then choose an action. The econometrician observes the joint distribution of public characteristics, group composition and actions.

## a. the game

The player set remains, as before,  $V$ , and type space remains the same as well, except that we will now dispense with  $v$ , the characteristic observable by those in the network but not by the econometrician. An individual now has two choices in the game. In the first stage, the individual chooses a location,  $a$  or  $b$ . In the second stage he plays the game described in Section 1 with everyone at his location, choosing  $\omega_i$  as before. At the end of the first stage, a group of people has formed at each location. We associate to each possible group  $g$  the sociocimatrices  $A^g$  and  $C^g$ . These are given exogenously, and are known *a priori* by the individuals in  $V$  and by the econometrician. Since this is just an extended example, we will simplify the discussion by choosing a particular  $A^g$  and  $C^g$ . We will assume for contextual effects that  $x$  is averaged equally over all individuals in *both* groups. We will assume for peer effects that individual  $i$  averages equally over all individuals other than himself in *his* own group. Peer effects are group-specific but contextual effects are not. This is a case where we would expect all parameters to be identified were there no endogeneity problem. We will also assume that the status-homophily term, with coefficient  $\beta$ , weights according to  $A^g$ .

The payoff function for individual  $i$  in a group with member set  $g \subset V$  depends only upon the characteristics and actions of members of  $g$ , and not on the location. The payoff function is<sup>15</sup>

$$\begin{aligned}
 u_i(g, \omega_i, \omega_{g/\{i\}}, \mathbf{x}_g) = & \\
 & \left( \gamma \mathbf{x}_i + \delta \sum_{j \in g} c_{ij}^g \mathbf{x}_j + v_i + \varepsilon_i \right) \omega_i - \frac{1}{2} \omega_i^2 - \\
 & \frac{\phi}{2} \left( \omega_i - \sum_{j \in g/\{i\}} a_{ij}^g \omega_j \right)^2 + \beta \left( \mathbf{x}_i - \sum_{j \in g/\{i\}} a_{ij}^g \mathbf{x}_j \right)^2
 \end{aligned} \tag{10}$$

---

<sup>15</sup> For a given vector  $z \in \mathbf{R}^{|V|}$  and group  $g$ ,  $z_g$  denotes the vector  $(z_i)_{i \in g}$ .



The payoff function is not yet completely specified, because it does not say what payoffs should be when  $i$  is in a group of one. We will assume that in this case,  $i$ 's social payoff is what they would receive were they in the other group. In other words, one cannot be a group of one.

In this payoff function there are two sources of homophily. If  $\beta$  is large relative to  $\phi$ , affiliation will be characteristic-based. This corresponds roughly with what is meant by status-homophily. If  $\beta$  is small relative to  $\phi$ , affiliation will be outcome based. Individuals who desire to behave in similar ways will be more likely to group together — roughly speaking, value homophily. Value-homophily as the source of endogeneity problems. Fix a value of  $\phi$ , and now perform the experiment of making  $\beta$  very large. For large enough  $\beta$  group membership is almost entirely determined by the direct effect of the publicly observable characteristics. The conditional probability distribution of group formation given  $x$  converges to a point mass as  $\beta$  diverges. Imagine the limit: For all but a measure-0 set of  $x$  values, the participation conditions defining group participation hold strictly. This means that a given group is stable under small perturbations of  $x$ . This is enough to recover  $B$ , and identification proceeds as in previous sections. It is important to note that not all sources of endogeneity lead to identification issues.

The action stage of the game requires a strategy profile  $f^g$  for every possible group that could form. A strategy profile for the first stage is an assignment of each individual to a location, a map  $\hat{\sigma}_i(x, \varepsilon_i) \rightarrow \{a, b\}$ . The assignment of individuals to locations maps each  $(x, \varepsilon_i)$  to a partition of  $V$  into two sets (one of which may be empty). We do not need to keep track of the locations, only the partition. Define  $\sigma(x, \varepsilon)$  to be the map to partitions defined by  $\hat{\sigma}$ . For any partition  $\{g, h\}$  the set  $\sigma(x, \cdot)^{-1}\{g, h\}$  is a product set in  $\mathbf{R}^{|V|}$  since each individual chooses a location seeing only his own  $\varepsilon_i$ . In the same manner, define  $\sigma_{-i}$  to be the induced partition on  $V / \{i\}$ , the partition of everyone other than  $i$ . Furthermore, given such a map  $\sigma$ , we can reverse the process

and construct a strategy profile  $\hat{\sigma}$  which would induce it. We will call the map  $\sigma$  an *assignment*, since it allocates individuals to groups.

The *interim payoff* to  $i$  for belonging to group  $g$  when all  $j \in g / \{i\}$  chosen according to the strategy profile  $f^g$  is

$$V_{ig}(\mathbf{x}, \varepsilon_i) = \sup_{\omega_i} \left( \gamma x_i + \delta \sum_{j \in g} c_{ij}^g x_j + \varepsilon_i \right) \omega_i - \frac{1}{2} \omega_i^2 - \frac{\phi}{2} E \left( \left( \omega_i - \sum_{j \in g / \{i\}} a_{ij}^g f_j^g(\mathbf{x}, \varepsilon_j) \right)^2 \middle| \sigma_{-i}(\mathbf{x}, \varepsilon_{-i}) = g / \{i\} \right) + \frac{\beta}{|g| - 1} \left( x_i - \sum_{j \in g / \{i\}} a_{ij}^g x_j \right)^2. \quad (11)$$

With interim payoffs in hand, we can define a perfect Bayes equilibrium of the two-stage game.

**Definition 3.** A profile  $\left( \sigma, (f^g)_{g \in P(V) / \emptyset} \right)$  is a perfect Bayes equilibrium if and only if

- i. Each  $f^g$  is a Bayes-Nash equilibrium of the second stage game for some conjectured  $(\mu_i^\varepsilon)_{i \in g}$ .
- ii. If  $g$  occurs with positive probability and  $i \in g$ , then  $\mu_i^\varepsilon = E(\varepsilon_i | \sigma(\mathbf{x}, \varepsilon) = g)$ .
- iii. For each  $x$  and  $g \in P(V) / \emptyset$  containing  $i$ , on the event  $\{\varepsilon : \sigma(\mathbf{x}, \varepsilon) = \{g, V / g\}\}$ ,  $V_{i,g}(\mathbf{x}, \varepsilon_i) \geq V_{i, V / g \cup \{i\}}(\mathbf{x}, \varepsilon_i)$

The first condition says that actual action choice in groups formed, and equilibrium conjectures for groups that do not form, are the Bayes-Nash equilibrium for that group for some conjectured assignment of individuals to groups. The second condition says that beliefs about the assignment have to be correct on the equilibrium path. The third

condition is a participation constraint. It says that no individual wants to change groups given the second-stage conjectures about group choices. We shall be computing equilibria which are symmetric in that two individuals in the same choice situation, that is, the same  $x$ ,  $\varepsilon_i$ , contextual effect, and expected peer effect, will choose the same way.

We will not prove existence here, but we will provide some characterization.<sup>16</sup> The characterization lemma states that if the unconditional (first-stage) expected value of the average choice of group  $g$  exceeds that of group  $V/g$ , and individual  $i$  with a given private type prefers  $g$  to  $V/g$ , then he will prefer  $g$  to  $V/g$  for all higher private types. That is, assignment rules in a stable equilibrium have a threshold property.

**Lemma 4.** If  $\rho > 0$  and, for individual  $i$  in equilibrium<sup>17</sup>

$$E\left(\sum_{j \in g} a_{ij}^g \omega_j\right) > E\left(\sum_{j \in V/g} a_{ij}^{V/g} \omega_j\right)$$

and if individual  $i$  with type  $\varepsilon_i$  weakly prefers  $g$  to  $h$ , then individual  $i$  with type  $\varepsilon'_i > \varepsilon_i$  will strictly prefer  $g$  to  $h$ . If  $\phi = 0$ , individual  $i$ 's group choice is determined solely by the direct homophily effect. He will join the group with characteristics most similar to his own. If  $\beta = 0$  as well, then each individual is indifferent over group choice.

The consequence of this Lemma is that for each  $g$ , the set of  $\varepsilon \in \mathbf{R}^M$  for which  $g$  forms is the product of intervals where each interval is either of the form  $[\varepsilon_i^g, \infty)$  or  $(-\infty, \varepsilon_i^g]$ . The source of the selection problem is similar to that which arises in discrete choice models. Selection is determined by a threshold in the space of private types,

---

<sup>16</sup>In general it is hard to prove the existence of a perfect Bayes equilibrium for games with a continuum of types. It is relatively straightforward to prove existence when the type space is finite, and also when there is no heterogeneity in publicly observable types (or no public observable type).

<sup>17</sup>If  $g = \{i\}$ , the sum over  $j \in g$  is replaced by  $g^c$ .

this threshold will change as we change individuals' observed characteristics, and so the mean of the private type of individual  $i$  conditional on being in group  $g$  will depend on the values of  $i$ 's characteristics and the characteristics of the other group members.

## b. identification

A structure for this game is a list with elements described in Definition 4 below. We have already described the sociomatrices above. We have added the homophily parameter  $\beta$ , and we have dispensed with  $\nu$ . The model  $M'$  of this section maintains properties iii and vi of Definition 1. Unfortunately, we have, so far, found no way to identify all of the parameters of the model without assuming *a priori* knowledge of  $\rho$ . The definition of structures and models for this section reflects all this:

**Definition 4.** A structure is a list  $\langle \gamma, \delta, \phi, \beta, C, A, \mu^\varepsilon, \rho \rangle$  satisfying Axioms 1 through 4.

Denote by  $M_{end}$  the set of structures satisfying the following additional properties:

- i.  $\rho^X$  is finitely exchangeable.
- ii.  $\rho^\varepsilon$  has a strictly positive density on  $\mathbf{R}^M$ .
- iii.  $\varepsilon_i$  is an independent and identically distributed sequence.
- iv. For all  $g$ ,  $A_{ij}^g = \frac{1}{|g|-1}$  if  $i$  and  $j \neq i$  are both in  $g$ , and 0 otherwise.
- v.  $c_{ij} = \frac{1}{|V|-1}$ .
- vi. One of  $\gamma$  and  $\delta$  is not 0.

- vii. One of  $\phi$  and  $\beta$  is greater than 0.

An observation in this model is a triple  $(\{g, V/g\}, x, \omega)$  where  $g$  and  $V/g$  are the two groups that form,  $x$  is the vector of characteristics of individuals in  $V$  and  $\omega$  is the vector of their actions. An observation is an *equilibrium outcome*. We will assume that econometricians have access to all data. That is, econometricians see who is in which group, and what each individual chooses. In other words, the econometrician sees a particular equilibrium assignment of individuals to groups, and the subsequent second stage equilibria for the (no more than) two groups that formed. Thus the identification question concerns probability distributions on triples of the form  $(\{g, V/g\}, x, \omega)$ .

A difficulty in addressing identification in strategic models is that the equilibrium need not be unique. We will assume that the econometrician knows which equilibrium describes the data. The state of the art on partial identification for games with multiple equilibria has not yet reached games of the kind we consider here. The following theorem summarizes identification in  $M_{end}$ .<sup>18</sup>

**Theorem 8.**

- i. If  $\mu^\varepsilon$  is known *a priori* by the econometrician, then  $\gamma$ ,  $\delta$ , and  $\phi$  are identified by the distribution of equilibrium outcomes.
- ii. If the conditional means  $E(\varepsilon|x, g)$  are known either *a priori* or identified by the distribution of equilibrium outcomes, and also known to be nonlinear, then  $\gamma$ ,  $\delta$ , and  $\phi$  are identified by the distribution of equilibrium outcomes.

---

<sup>18</sup>It leaves out one case: if  $\rho = 0$ , then the most that can be said about  $\beta$  is whether it is 0 or positive, because in this case group assignment probabilities are independent of the (nonzero) magnitudes.

- iii. If  $\rho$  is known *a priori* and  $\phi > 0$ , then  $\beta$  is identified.

The difficulty with identification is recovering the matrix  $B(m)$ , where all parameters except  $\beta$  are hiding. When the network is exogenous, varying the  $x$ 's a bit and seeing what happens uncovers the linear relation between  $x$  and  $\omega$ , in other words,  $B(m)$ , and identification proceeds from there. When networks are endogenous, varying  $x$  changes the participation constraints. The conditional mean of  $\varepsilon_i$  given  $x$  in each group move with  $x$ . Perturbations in  $x$  perturb terms that are constants with exogenous networks, so picking out  $B(m)$  becomes a nontrivial task. When the unconditional  $E(\varepsilon)$  is known, we bypass  $B(m)$  and go directly for the parameter values. When the conditional means  $E(\varepsilon|x,g)$  are known *a priori* or can be estimated by other means, we can subtract off their contribution to  $\omega$  and recover  $B(m)$ . Heckman's work on self-selection provides one path into estimating  $E(\varepsilon|x,g)$  when they are in fact identified, a question which we explore in the next section, although a comprehensive treatment is beyond the scope of this paper.

## ii. econometric approaches to identification with endogenous networks

Heckman's early classic work (Heckman (1979)) has evolved into the control function approach to self-selection (e.g. Heckman and Robb (1985, Section 3.4, 1986)<sup>19</sup>. Formally, we define a control function by the requirement that

$$s_i \propto E(\varepsilon_i|x,g) \tag{12}$$

so that for some  $\theta$ ,

---

<sup>19</sup> See Navarro (2008) for a recent overview.

$$\varepsilon_i = \theta s_i + \zeta_i \quad (13)$$

where

$$E(\zeta_i | \mathbf{x}, \mathbf{g}) = 0. \quad (14)$$

Equation (12) implies that when agent  $i$  forms expectations of  $\omega_{-i}$ , this expectation will differ from the case when the network is exogenous as modeled in Section 4. However, the information set on which the agent conditions is the same as in the original model. Hence the control function approach amounts to analyzing the equation

$$\omega_i = \frac{\gamma}{1+\phi} \mathbf{x}_i + \frac{\delta}{1+\phi} \sum_j \mathbf{c}_{ij}^g \mathbf{x}_j + \frac{\phi}{1+\phi} \sum_j E(\mathbf{a}_{ij}^g \omega_j | \mathbf{x}, \mathbf{g}) + \pi s_i + \frac{\zeta_i}{1+\phi} \quad (15)$$

It is evident that the presence of  $s_i$  as a regressor in (15) converts the equation into one in which the regressors are orthogonal to the regression residual. Of course, it will be necessary for (12) to be nonlinear in order to avoid linear dependence on the other regressors in the equation. This is true outside of special cases for group formation. Note as well that the variables  $s_i$  are not associated with contextual effects in (15). Hence, when they are nonzero, it is the case that  $E(\omega_j | \mathbf{x}, \mathbf{g})$  is no longer linearly dependent on the set of  $\mathbf{x}_j$ 's,  $j \in g$  which is the source of the reflection problem when the sociomatrices produce equations of the form (7). This is an example in which endogenous network formation produces identification when exogenous network formation would not; see Brock and Durlauf (2001b, 2006) for more discussion.<sup>20</sup>

---

<sup>20</sup>Our claim that network formation can enhance identification of network effects is paralleled in Heckman's research which shows that self-selection such as failure to comply to a treatment contains economically relevant information. The counterintuitive idea that noncompliance can aid identification is another instance of the general

We have said nothing about how to construct the control function or whether they even exist. It is now understood that control function may not exist in certain contexts (Blundell and Matzkin (2010)). Our goal is simply to establish how one could in principle use endogenous network formation to facilitate identification of our general social networks model, so long as the control function approach can be implemented. Notice this is exactly the import of Theorem 8.ii above.

The idea that self-selection can aid in identification of social effects via control functions was first shown in Brock and Durlauf (2001b); in this case the  $s_i$ 's turn out to be proportional to the Heckman  $\lambda$ 's from Heckman's early work on correction for selection bias (Heckman (1979)). Brock and Durlauf (2006) provide a more general treatment when agents select into cliques and weights are required to be equal; for this environment the  $s_i$ 's correspond to the generalization of the original Heckman selection correction proposed by Lee (1983). These papers show that the set of social networks models for which one can construct control functions is not empty. Ioannides and Zabel's (2008) housing market study shows that there exist contexts in which the control function approach can be empirically implemented. We leave the question of the extension of the approach outlined here to general networks to future research. Our main message is that if the control approach is implementable, then subject to standard conditions on regressors, identification can be achieved for endogenous networks.

This all said, control functions are not a panacea. Endogeneity has a particular source in social interaction models. Network formation and action on the network are the two parts of a multistage game. Considerations from the underlying game suggest important limits on the way this procedure can be conducted. It is important that equation (14) not be interpreted as a behavioral equation without consideration of the first stage of the game. This is evident when one considers how the validity of equation (12) played a critical role in the analysis.

---

principle that choices that have traditionally been regarded as nuisances in fact encode information that can be exploited by the econometrician. See Heckman (1992) and Heckman and Smith (1995,1997) for formalization of the principle for the case of noncompliance.



To see the import of this argument, suppose it is the case that there is additional information  $z$  that affects the choice of networks in the first stage of the game but has no effect on the payoffs associated with the choices  $\omega_i$  in the second stage of the game. At first glance, one might believe that  $z$  represents a set of instruments available for overcoming endogeneity of the group choices, that may be used to overcome the correlation of regression errors and regressors in an equation such as equation (3) when networks are endogenous. One might even consider control functions of the form  $s_i \propto E(\varepsilon_i | x, z)$  as candidates for instruments and go so far as to conclude that the  $\frac{\phi}{1+\phi} a_{ij}^g$  terms can be identified using these instruments, and so resolve the problem of an unobserved  $A$  matrix. However, such mechanical reasoning would not be appropriate. The existence of  $z$  as a set of factors that determine group selection will affect the form of the second stage equation for choices within a network if they are available to the individuals in the network. In other words, equation (3) will not generally hold in the presence of  $z$  in the first stage and so (14) would be misspecified. Of course, if it were the case that the  $s$  vector constitutes data observable by individuals only after they have chosen their group, there is no problem. And finally, there is no reason in principle this could be data available to the econometrician but not the individuals in the network. In summary, the structural model of network formation is needed to provide guidance for the choice of appropriate instruments, guidance that would not be readily apparent were one to simply consider equation in isolation.

## 7. Conclusions

In this paper, we have provided a theoretical and econometric characterization of linear social interactions models. These models represent the workhorse of much of the current empirical research in social economics. Our analysis provides both a clear description of the behavioral assumptions needed to employ these models as well as the conditions under which the primitive utility parameters that characterize individual and social influences may be recovered. The results indicate the importance of prior

information on social network structure and highlight the importance of data collection as an integral part of efforts to identify economically interesting phenomena. At the same time, our analysis shows that identification is not only a function of what data are available, but of the features of the social networks themselves. As such, they illustrate a range of cases when identification will and will not hold. An important feature of our results is that we are able to specify how different aspects of socioeconomic environment can be identified, depending on the nature of a researcher's *a priori* information. Hence, we find that it is possible to identify the intensity of peer group effects even if the identities and averaging rule of an individual is unknown.

In terms of future research, we see two important directions. First, our analysis has explored the polar cases where the social networks that embed individuals are and are not observed. The question of identification in the presence of partial observability has yet to be systematically studied. We have referred to one form of partial observability, namely knowledge of the zeroes in the relevant sociomatrices, in our discussion of the Add Health data set. These work as exclusion restrictions from the vantage point of classical simultaneous equations theory, and as such can provide identification under partial observability. But one can, for example imagine distinct questions involving identification when only a subset of network members are observed. While this problem often arises, its implications for identification have yet to be assessed. Further, it would seem natural, when surveys can only obtain information from a subset of a population, that survey design should be constructed in order to facilitate identification. Second, our analysis has not addressed the question of what can be uncovered when a network is evolving. Our analysis has taken the network as fixed. However, the fact that different networks may or may not be identified suggests that networks may evolve through periods in which behavioral parameters are and are not identified. For stochastic network formation processes, this leads to the interesting question of the probability that the network passes through a period when identification is possible.

Further, while we have addressed the question of how our identification results are affected by network endogeneity, we have not addressed how this endogeneity can, when explicitly modeled, facilitate identification, although we are hinted at this

possibility. For example, if network membership is associated with prices, then prices can help to uncover social effects, as demonstrated in recent advances in the econometrics of hedonic models (Ekeland, Heckman, and Nesheim (2004), Nesheim (2002)). One of the major themes in James Heckman's research that endogeneity is not so much a nuisance to empirical work, but rather an additional behavior that needs to be modeled. So our last suggestion is nothing more but an acknowledgement of the importance of this particular Heckman insight to future social networks research.

## Technical Appendix

### Section 2.

#### Proof of Theorem 1.

$$E(u_i(\omega_i, \omega_{-i}) | \mathbf{x}, \nu, \varepsilon_i) = \left( \gamma \mathbf{x}_i + \delta \sum_j \mathbf{c}_{ij} \mathbf{x}_j + \nu_i + \varepsilon_i \right) \omega_i - \frac{1}{2} \omega_i^2 - \frac{\phi}{2} E \left( \left( \sum_j \mathbf{a}_{ij} \omega_j - \omega_i \right)^2 | \mathbf{x}, \nu, \varepsilon_i \right).$$

Let  $\psi_i = \gamma \mathbf{x}_i + \delta \sum_j \mathbf{c}_{ij} \mathbf{x}_j + \nu_i + \mu_i^\varepsilon$  and let  $\varepsilon_i^{\text{dev}}$  denote the deviation of  $\varepsilon_i$  from its mean.

The common knowledge assumption implies that the vector  $\mu^\varepsilon$  is known to all network members, so it is only the deviation from the mean that is private. Rewriting,

$$E(u_i(\omega_i, \omega_{-i}) | \mathbf{x}, \nu, \varepsilon_i^{\text{dev}}) = (\psi_i + \varepsilon_i^{\text{dev}}) \omega_i - \frac{1}{2} \omega_i^2 - \frac{\phi}{2} E \left( \left( \sum_j \mathbf{a}_{ij} \omega_j - \omega_i \right)^2 | \mathbf{x}, \nu, \varepsilon_i^{\text{dev}} \right).$$

The first-order conditions are

$$\psi_i + \varepsilon_i^{\text{dev}} - \omega_i - \phi \left( \omega_i - \sum_j \mathbf{a}_{ij} E(\omega_j | \mathbf{x}, \nu, \varepsilon_i^{\text{dev}}) \right) = 0$$

and so

$$\omega_i = \frac{1}{1+\phi} \psi_i + \frac{\phi}{1+\phi} \sum_j \mathbf{a}_{ij} E(\omega_j | \mathbf{x}, \nu, \varepsilon_i) + \frac{1}{1+\phi} \varepsilon_i^{\text{dev}}$$

which implies

$$E(\omega_i | x, \nu) = \frac{1}{1+\phi} \psi_i + \frac{\phi}{1+\phi} \sum_j a_{ij} E(\omega_j | x, \nu).$$

Thus

$$E(\omega | x, \nu) = \frac{1}{1+\phi} \left( I - \frac{\phi}{1+\phi} A \right)^{-1} \psi.$$

This means that

$$\omega = \frac{1}{1+\phi} \psi + \frac{\phi}{1+\phi} A \frac{1}{1+\phi} \left( I - \frac{\phi}{1+\phi} A \right)^{-1} \psi + \frac{1}{1+\phi} \varepsilon^{dev}$$

Rearranging terms, the set of choices follows

$$\begin{aligned} \omega &= \frac{1}{1+\phi} \left( I - \frac{\phi}{1+\phi} A \right)^{-1} \psi + \frac{1}{1+\phi} \varepsilon^{dev} = \\ &= \frac{1}{1+\phi} \left( I - \frac{\phi}{1+\phi} A \right)^{-1} \left( (\gamma I + \delta C) x + \nu + \mu^\varepsilon \right) + \frac{1}{1+\phi} \varepsilon^{dev} \end{aligned}$$

which is therefore an equilibrium.

Uniqueness, of equilibrium is proven by that the first order conditions define a contraction map on the space of strategy profiles topologized with the product  $L_2$  norm.

This space is not empty, and if  $f$  is in this space,  $E|f_i(\psi, \varepsilon_i)|^2 < \infty$ . Define the operator

$$(Tf)_i(\psi, \varepsilon_i) = \frac{1}{1+\phi} (\psi + \varepsilon_i) + \frac{\phi}{1+\phi} E(a_i \cdot f(\psi, \varepsilon)).$$

A fixed point of this map is a strategy profile that satisfies the first-order condition for every agent, and hence is a Bayes-Nash equilibrium. A straightforward computation shows that  $T$  is a contraction mapping, and so its fixed point is unique.  $\square$

#### Section 4.

#### Proof of Lemma 2.

From equation (4) and the hypothesis of the lemma it follows that

$$(1 + \phi)B - \phi AB = \gamma I + \delta C \quad (16)$$

for any  $(\phi, \delta, \gamma) \in B^{-1}(B)$ . Choose an  $(i, j)$  pair satisfying the hypothesis of the lemma.

The right hand side of (16) is 0, and so

$$\frac{\phi}{1 + \phi} = \frac{b_{ij}}{[AB]_{ij}}.$$

(Note that the denominator on the right cannot be 0, or else from (16),  $\phi = -1$  which is satisfied by no model in  $M$ .) Thus if  $B(m') = B(m'') = B$ , then  $\phi' = \phi''$ . From property ii of the Definition 1 for  $M$ , there is an  $(i, j)$  pair with  $i \neq j$  such that  $c_{ij} \neq 0$ , and so from (16),  $\delta' = \delta''$ , and the equation for any diagonal pair implies that  $\gamma' = \gamma''$ . From these equalities it follows that  $\mu^{v'} + \mu^{\varepsilon'} = \mu^{v''} + \mu^{\varepsilon''}$ .  $\square$

#### Proof of Theorem 2.

Suppose that  $B(m') = B(m'') = B$ . We can write

$$B = \frac{1}{1+\phi'} \left( I - \frac{\phi'}{1+\phi'} A \right)^{-1} (\gamma' I + \delta' C) =$$

$$\frac{1}{1+\phi'} \left( I + \frac{\phi'}{1+\phi'} A + \left( \frac{\phi'}{1+\phi'} \right)^2 A^2 + \dots \right) (\gamma' I + \delta' C)$$

Since the peer-effects network is connected, some power of  $A$  is strictly positive. Suppose that  $m'$  is such that  $\phi' > 0$ . Then  $(1-\phi)^{-1} (I - (\phi'/(1+\phi')) A)^{-1}$  is strictly positive. The assumption on  $C$  of no zero columns ensures that  $A^n C \gg 0$  if  $A^n \gg 0$ .<sup>21</sup> Thus  $(I - \phi'/(1+\phi') A)^{-1} C$  is strictly positive. Choose an  $i$  and  $j$  for which  $c_{ij} = 0$ . The set of all  $\gamma, \delta$  pairs that can make  $b_{ij} = 0$  is a 1-dimensional line in  $\mathbf{R}^2$ . If  $(\gamma'', \delta'')$  is not on this line, then according to Lemma 1,  $m'' = m'$ . The set of  $(\phi, \gamma, \delta)$  triples for which  $b_{ij} = 0$  is a closed, two-dimensional semi-algebraic set<sup>22</sup>; the set of models in  $M^{AC}$  with parameters outside this set is generic. Thus we have generic identification of  $\phi$ ,  $\gamma$  and  $\delta$  for any  $B = B(m)$  with  $\phi > 0$ . If  $\phi' = 0$ , then generically  $\phi'' = 0$  (a consequence of the preceding argument). In this case it is straightforward to see that  $\gamma' = \gamma''$  and  $\delta' = \delta''$ . Finally, in either case, if everything else is equal, it follows that  $\mu^{v'} + \mu^{\varepsilon'} = \mu^{v''} + \mu^{\varepsilon''}$ .  $\square$

### Proof of Theorem 3.

Choose  $i$  and  $j$  in  $V_1^C \cap V^A$  and  $V_2^C \cap V^A$ , respectively. The matrix  $(1/(1+\phi))(I - \phi/(1+\phi) A)^{-1}$  is block diagonal, with strictly positive blocks corresponding to the different components of  $A$ . Therefore

<sup>21</sup>  $X \gg 0$  means that every element of the matrix  $X$  is strictly positive.

<sup>22</sup> See Bochnak, Coste, and Roy (1998) for a comprehensive overview of semi-algebraic sets.

$$b_{ij} \geq \frac{1}{1+\phi} \sum_{k \in V_2^C \cap V^A} \left[ \left( I - \frac{1}{1+\phi} A \right)^{-1} \right]_{ik} c_{kj}.$$

Transitivity implies that the component  $V_2^C$  is a clique—completely connected—and so this sum is positive. But  $c_{ij} = 0$  by assumption ( $i$  is not in  $V_2^C$ ). Thus the hypothesis of Lemma 2 is satisfied.  $\square$

### Proof of Lemma 3.

Identification holds if and only if, for each matrix  $B$ ,  $B^{-1}(B)$  generically produces unique parameters. So suppose  $B(m') = B(m'') = B$ . Then

$$\left( I - \frac{\phi'}{1+\phi'} A \right)^{-1} (\gamma' I + \delta' C) = \left( I - \frac{\phi''}{1+\phi''} A \right)^{-1} (\gamma'' I + \delta'' C),$$

so

$$\left( I - \frac{\phi''}{1+\phi''} A \right) (\gamma' I + \delta' C) = \left( I - \frac{\phi'}{1+\phi'} A \right) (\gamma'' I + \delta'' C)$$

since the matrices commute, and so

$$(\gamma' - \gamma'') I + (\delta' - \delta'') C + \left( \frac{\phi'}{1+\phi'} \gamma'' - \frac{\phi''}{1+\phi''} \gamma' \right) A + \left( \frac{\phi'}{1+\phi'} \delta'' - \frac{\phi''}{1+\phi''} \delta' \right) AC = 0. \quad (17)$$

We specialize equation (17) to the various cases.

1. If the matrices are linearly independent, then the coefficients of the four matrices must each be 0. Thus  $\gamma' = \gamma''$  and  $\delta' = \delta''$ . One of  $\gamma'$  and  $\delta'$  is not 0, so at least one of



the last two terms implies that  $\phi' = \phi''$ . Conversely, suppose that the matrices are linearly dependent, and suppose that that

$$aI + bC + cA + dAC = 0$$

for  $a$ ,  $b$ ,  $c$  and  $d$  not all 0. We will construct two (in fact, many) models  $m'$  and  $m''$  which give rise to the same  $B$ . For any  $\phi$ , let  $r = \phi/(1+\phi)$ . If two models  $m'$  and  $m''$  cannot be distinguished, the following equations must be satisfied:

$$\begin{aligned}\gamma'' &= \gamma' - a \\ \delta'' &= \delta' - b \\ c &= r'\gamma'' - r''\gamma' \\ d &= r'\delta'' - r''\delta'.\end{aligned}$$

Choose any  $r' \neq r''$  in  $[0, 1)$ . Substitute the first two equations into the last two to get

$$\begin{aligned}c &= (r' - r'')\gamma' - r'a \\ d &= (r' - r'')\delta' - r'b,\end{aligned}$$

and so solving for  $\gamma'$  and  $\delta'$  and working backwards gives parameters  $(\gamma', \delta', \phi')$  and  $(\gamma'', \delta'', \phi'')$  for the two structures  $m'$  and  $m''$ . (We have the requirements that  $\gamma'\delta', \gamma''\delta'' \neq 0$ . This will clearly be satisfied for generic choices of  $r'$  and  $r''$ .) To complete the description, choose the same prior distribution  $\rho$  for both models satisfying the requirements of Axioms 1-4 and conditions i - iii of the Definition 1 for  $M$ .

2. Substitute  $C$  for  $AC$  and regroup the terms of (17) and suppose again that  $B(m') = B(m'') = B$ . If the matrices are independent, then  $\gamma' = \gamma''$  without further assumptions. Since  $\gamma' \neq 0$  independence implies that  $\phi' = \phi''$ . If  $\phi' \neq 0$ , then  $\delta' = \delta''$

(since  $\phi'/(1+\phi'') < 1$ ). If  $\delta' = 0$ ,  $B(m') = B(m'')$  implies that  $\gamma'I + \delta'C = \gamma''I + \delta''C$ , and identification holds.

3. The case  $A = C$  is proved in Bramoullé, Djebbari and Fortin (2009).  $\square$

### Proof of Corollary 1.

This is case 2 of Lemma 2,  $AC = C$ . The three matrices  $I$ ,  $A$  and  $C$  are dependent if and only if  $A$  is a linear combination of  $I$  and  $C$ . But any such linear combination has to have identical off-diagonal elements.  $\square$

### Section 5.

#### Proof of Theorem 5.

Proof. Suppose  $E(\omega | x, m') = E(\omega | x, m'')$  on some open subset of  $\mathbf{R}$ . Then the strategy profiles for structures  $m'$  and  $m''$  are described by the same matrix  $B$ . Suppose too that  $\gamma \neq 0$ .

$$\frac{\gamma'}{1+\phi'} \left( I - \frac{\phi'}{1+\phi'} A' \right)^{-1} = \frac{\gamma''}{1+\phi''} \left( I - \frac{\phi''}{1+\phi''} A'' \right)^{-1} \quad (18)$$

That is,

$$\frac{\gamma'}{1+\phi'} \left( I - \frac{\phi''}{1+\phi''} A'' \right) = \frac{\gamma''}{1+\phi''} \left( I - \frac{\phi'}{1+\phi'} A' \right)$$

Since the diagonal elements of  $A'$  and  $A''$  are 0, it follows that

$$\frac{\gamma'}{1+\phi'} = \frac{\gamma''}{1+\phi''}$$

Since  $\gamma'$  is non-zero, the corresponding  $\gamma''$  cannot equal 0, and it follows that

$$I - \frac{\phi'}{1+\phi'} A' = I - \frac{\phi''}{1+\phi''} A'',$$

which immediately implies

$$\frac{\phi'}{1+\phi'} A' = \frac{\phi''}{1+\phi''} A''.$$

From the observation that the rows of both  $A'$  and  $A''$  sum to 1, one may conclude that  $\phi' = \phi''$ , and therefore  $\gamma' = \gamma''$ , and  $A' = A''$ . Finally from these facts it follows that  $\mu^v + \mu^\varepsilon$  is the same in both models.  $\square$

### Proof of Theorem 6.

Let  $B \in B\left(M_{(\gamma I + \delta C)}^C\right)$ . It follows from equation (4) that  $B$  has constant row sums, which we will call  $b$ . A computation show that  $b = \gamma + \delta$ . Rewriting (4),

$$(1+\phi)I - \phi A = (b-\delta)B^{-1} + \delta CB^{-1} \tag{19}$$

Consider the right-hand side as a function of  $\delta$ . Since  $A$  has 0's on the diagonal, it follows that there must be at least one value of  $\delta$  for which all the diagonal elements of the matrix on the right are equal. Since the right-hand side is linear in  $\delta$ , equality of the diagonal elements is true for either one  $\delta$  or all  $\delta$ . Choose any  $\delta'$  for which the

diagonal elements are equal. Then  $\phi = 1 - \left[ (b - \delta')B^{-1} + \delta CB^{-1} \right]_{11}$ , proving the first claim of the theorem. Next, it is easy to verify that for generic  $C$ , the  $\delta'$  which makes the diagonal elements identical is unique. This identifies  $\delta$ , and then  $\gamma = b - \delta$ .  $\square$

### Proof of Theorem 7.

A calculation shows that that  $D_{\gamma, A, C}B$  (The derivative map of  $B$  with respect to  $\gamma$ ,  $A$ , and  $C$ ) is surjective for all models in the interior of  $M$ , and so the implicit function theorem implies that if  $B(\phi', \gamma', \delta', A', C') = B$ , and  $(\phi'', \delta'')$  is sufficiently near to  $(\phi', \delta')$ , there are parameters  $\gamma''$ ,  $A''$  and  $C''$  such that  $B(\phi'', \gamma'', \delta'', A'', C'') = B$ .  $\square$

## Section 6

### Proof of Lemma 4.

The direct homophily effect, scaled by  $\beta$ , has no effect on individuals' preferences. Let  $\tilde{g}_1$  denote the set-valued random variable whose values are the members of group 1 realized at the end of the first stage. The value of being in a group 1 conditional on  $\tilde{g} = g$  is

$$V_{ig} = \frac{1}{2(1+\phi)} \left( \phi^2 E(\bar{\omega}_{-i} | g)^2 - \rho(1+\rho) E(\bar{\omega}_{-i}^2 | g) + 2E(\bar{\omega}_{-i} | g)(x_i + \varepsilon_i) + (x_i + \varepsilon_i)^2 \right) - \beta \sum_{j \in g} (x_j - x_i)^2$$

where  $\bar{\omega}_{-i}$  is the average choice of members of  $g$ . (This comes from the first order conditions, and substituting back.) The utility difference between  $g$  and  $h = V/g$  is

$$\begin{aligned} \Delta V_{gh} = & \frac{\rho}{2(1+\rho)} \left( 2(E(\bar{\omega}_{-i} | g)) - E(\bar{\omega}_{-i} | h) \varepsilon_i \right) \\ & - 2 \left( E(\bar{\omega}_{-i}^2 | g) - E(\bar{\omega}_{-i}^2 | h) \right) - \rho \left( \text{var}(\bar{\omega}_{-i} | g) - \text{var}(\bar{\omega}_{-i} | h) \right) \\ & - \beta \left( \sum_{j \in g} (x_j - x_i)^2 - \sum_{j \in h} (x_j - x_i)^2 \right) \end{aligned}$$

Now take expectations over  $\tilde{g}$  to see that if  $E(\bar{\omega}_{-i} | g) - E(\bar{\omega}_{-i} | h) > 0$ , then  $\Delta V_{gh}$  is increasing in  $\varepsilon_i$ , and so  $g$  is preferred to  $V/g$  whenever  $\varepsilon_i$  is large enough.  $\square$

### Proof of Theorem 8.

We prove this by solving the game for a several different choices of  $x$ , and using the answers together to make inferences about parameter values. First, consider identical  $x_i$ 's for every individual, say  $x_i \equiv 1$ . It is straightforward to show that there is a common cutoff  $\varepsilon^*$  such that individuals with  $\varepsilon_i > \varepsilon^*$  go to, say, location  $a$ , and the remainder go to  $b$ . For the individual who choose location  $z$ ,

$$\omega_i = \gamma + \delta + \frac{\phi}{1+\phi} E_z(\varepsilon_i) + \frac{1}{1+\phi} \varepsilon_i$$

where  $E_z(\varepsilon_i)$  is the expected value of  $\varepsilon_i$  conditional on the location, that is, either above or below  $\varepsilon^*$ . (Recall that the  $\varepsilon_j$  are all iid.) Compute the expected value of  $\omega_i$  at each location:

$$E_z(\omega_i) = \gamma + \delta + E_z(\varepsilon_i).$$

The expectations on the left are known to the econometrician. The probability of appearing at  $a$  is  $\Pr(\varepsilon_i \geq \varepsilon^*)$ . Thus we have a third equation

$$\Pr(\varepsilon_i \geq \varepsilon^*)E_a(\varepsilon_i) + \Pr(\varepsilon_i < \varepsilon^*)E_b(\varepsilon_i) = E(\varepsilon_i).$$

If the econometrician knows  $E(\varepsilon_i)$ , he can compute the three unknowns; the two conditional expectations and  $\gamma + \delta$ .

Next we examine a second-stage game in which one individual, say individual 1, has  $x_1 = k$ , and the remaining  $x_i = 1$ . Again, there are thresholds  $\varepsilon_1^*$  and  $\varepsilon_2^*$  for the players with  $x_i = 1$  and 2, respectively, and we suppose that those above the threshold go to  $a$  while those below go to location  $b$ . An equilibrium computation shows the following: If a group  $g$  which includes individual 1 assembles at  $a$ , then the second stage equilibrium has the property that for any person  $i \neq 1$  in  $g$ ,

$$E_a(\omega_1 | x_1 = k) = -E_a(\omega_i | x_1 = k) = \frac{|g| - 1}{|g| - 1 + r} (E_a(\varepsilon_1 | x_1 = k) - E_a(\varepsilon_i | x_1 = k) - (k - 1)\gamma)(1 - r),$$

where  $r = \phi / (1 + \phi)$ , and that  $E_a(\omega_1 | x_1 = k) > E_a(\omega_i | x_1 = k)$ . The ratio of this difference for two differently-sized groups with the same  $k$  determines  $\phi$ .

Furthermore, if  $g^c$  has at least two members, we also know that for this group,  $E_b(\omega_i | x_1 = k) = \gamma + \delta + E_b(\varepsilon_i | x_1 = k)$ . Since  $\gamma + \delta$  has already been identified and since  $E_b(\omega_i | x_1 = k)$  is observable,  $E_b(\varepsilon_i | x_1 = k)$  can be computed. This is the same for all individuals with  $x = 1$ , and is group-independent. The probability that individual  $i$  locates at  $b$  is observable, and  $E(\varepsilon_i)$  is *a priori* knowledge, so  $E_a(\varepsilon_i | x_1 = k)$  is identified. This gives the equation

$$E_a(\varepsilon_1 | x_1 = k) - k\gamma = z_a^k$$

where  $z_k$  can be computed from what is observable and the *a priori* knowledge of  $E(\varepsilon_i)$ . Now repeat the construction with individual 1 at location  $b$  to get

$$E_b(\varepsilon_1 | x_1 = k) - k\gamma = z_b^k$$

Since the probability of individual 1 appearing at each location is known, expect over the location to derive

$$E(\varepsilon_1) - k\gamma = \Pr(\varepsilon_1 \geq \varepsilon_1^*) z_a^k + \Pr(\varepsilon_1 < \varepsilon_1^*) z_b^k,$$

and so  $\gamma$  can be computed. Knowing  $\gamma$  gives  $\delta$ , which proves the claim for  $M_{end}$  with known  $E(\varepsilon_1)$ .

If the conditional expectations are known, take any group  $g$  with more than 1 person that forms with positive probability given  $x$ . Then  $E(\omega | x, g) - E(\varepsilon | x, g)$  is linear in  $x$ . The probabilities of entry into  $g$  are continuous in  $x$ , so this difference in fact is well-defined on an open set around  $x$ . From this infer  $g$  to get  $B(m)$ . The result now follows from the proof of theorem 2.

The parameter  $\beta$  has no role in the second stage game. It determines only the probabilities of group formation. Consequently  $\beta$  must be identified off the participation constraint, that the ex-ante expected value of going to location  $a$  is at least that of going to  $b$  for those who chose to go to  $a$ , and so forth. The threshold  $\varepsilon$  in any game is determined by the equality of the expectation of the interim payoffs over which groups will form at  $a$  and  $b$  given the second-stage equilibrium strategies. Consider then, a situation with heterogeneous  $x_i$ 's. On the one hand,  $\varepsilon^*$  is known given *a priori*

knowledge of  $\rho$ , because the probabilities of a given individual appearing at either location in equilibrium are known, and the distribution function of  $\varepsilon_i$  is strictly increasing. Then the equilibrium condition defining  $\varepsilon^*$ , that when  $\varepsilon_i = \varepsilon^*$  the individual is indifferent (ex-ante) between  $a$  and  $b$ , identifies  $\beta$ . The prior belief  $\rho$  is needed not just to determine  $\varepsilon^*$ , but also because the participation constraints involve differences in variances of the  $\varepsilon_i$  conditional on location, and there are not enough equations to pin these down even given *a priori* knowledge of the unconditional variance of the  $\varepsilon_i$ .  $\square$



## Bibliography

- Aizer, A. and J. Currie. 2004. "Networks or Neighborhoods? Correlations in the Use of Publicly Funded Maternity Care in California." *Journal of Public Economics* 88 : 2573-2585.
- Ballester, C., A. Calvó-Armengol, and Y. Zenou. 2010. "Delinquent Networks." *Journal of the European Economic Association* 8: 34-61.
- Bandiera, O. and I. Rasul. 2006. "Social Networks and Technology Adoption in Northern Mozambique." *The Economic Journal*, 166: 869-902.
- Bayer, P., S. Ross, and G. Topa. 2008. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes." *Journal of Political Economy* 116: 1150-1196.
- Bertrand, M., E. Luttmer, and S. Mullainathan. 2000. "Network Effects and Welfare Cultures." *Quarterly Journal of Economics* 115: 1019-1055.
- Becker, G. and K. Murphy. 2001. *Social Economics*. Cambridge: Harvard University Press.
- Benhabib, J., A. Bisin, and M. Jackson, eds. 2010. *Handbook of Social Economics*. Amsterdam: North Holland.
- Blume, L., W. Brock, S. Durlauf and Y. Ioannides. 2010. "Identification of Social Interactions." In *Handbook of Social Economics*, J. Benhabib, A. Bisin, and M. Jackson, eds. Amsterdam: North Holland.
- Blundell, R. and R. Matzkin. 2010. "Conditions for the Existence of Control Functions in Nonseparable Simultaneous Equations Models." Mimeo, University College London.
- Bobonis, G. and F. Finan. 2009. "Neighborhood Peer Effects in Secondary School Enrollment Decisions." *Review of Economics and Statistics* 91: 695-716.
- Bochnak, J., M. Coste, and M.-F. Roy. 1998. *Real Semi-Algebraic Geometry*. Heidelberg: Springer-Verlag.
- Bramoullé Y., H. Djebbari, and B. Fortin. 2009. "Identification of Peer Effects through Social Networks." *Journal of Econometrics* 150: 41–55.
- Brock, W. and S. Durlauf. 2001a. "Discrete Choice with Social Interactions." *Review of Economic Studies* 68: 235-60.
- Brock, W. and S. Durlauf. 2001b. "Interaction-Based Models." In *Handbook of Econometrics*, Vol. 5, J. Heckman and E. Leamer, eds. Amsterdam: North-Holland

Brock, W. and S. Durlauf. 2006. "Multinomial Choice with Social Interactions." In *The Economy as an Evolving Complex System*, vol. 3, L. Blume and S. Durlauf, eds. New York: Oxford University Press.

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

Christakis, N. and J. Fowler, J. 2007. "The Spread of Obesity in a Large Social Network over 32 Years." *New England Journal of Medicine* 357: 370-379.

Cohen-Cole, E. and J. Fletcher. 2008. "Is Obesity Contagious? Social Networks vs. Environmental Factors in the Obesity Epidemic." *Journal of Health Economics* 27: 1382-1387.

Conley, T. and G. Topa. 2002. "Socio-Economic Distance and Spatial Patterns in Unemployment." *Journal of Applied Econometrics* 17: 303-327.

Conley, T. and G. Topa. 2003. "Identification of Local Interaction Models with Imperfect Location Data." *Journal of Applied Econometrics* 18: 605-618.

Conley, T. and C. Udry. 2010. "Learning About a New Technology: Pineapple in Ghana." *American Economic Review* 100: 35-69.

Cooley, J. 2008. "Desegregation and the Achievement Gap: Do Diverse Peers Help?" Unpublished manuscript. Department of Economics, University of Wisconsin at Madison.

De Giorgio, G., M. Pellizzari, and S. Redaelli. 2010. "Identification of Social Interactions through Partially Overlapping Peer Groups." *American Economic Journal: Applied Economics* 2: 241-275.

Dobrushin, R. 1965. "Existence of a Phase Transition in Two and Three Dimensional Ising Models." *Theory of Probability and its Applications*. 10: 193-213.

Drton, M., R. Foygel, and S. Sullivant. 2011. "Global Identifiability of Linear Structural Equation Models." *Annals of Statistics* 39: 865-886.

Durlauf, S. 2002. "On the Empirics of Social Capital." *Economic Journal* 112: 459-479.

Ekeland, I., J. Heckman and L. Nesheim. 2004. "Identification and Estimation of Hedonic Models," *Journal of Political Economy* 112: S60-S109.

Fisher, F. 1966. *The Identification Problem in Econometrics*. New York: McGraw-Hill.

- Glaeser, E., B. Sacerdote, and J. Scheinkman. 1996. "Crime and Social Interactions." *Quarterly Journal of Economics* 111: 507-548.
- Glaeser, E., B. Sacerdote, and J. Scheinkman. 2003. "The Social Multiplier." *Journal of the European Economic Association* 1: 345-353.
- Goyal, S. 2009. *Connections*. Princeton: Princeton University Press.
- Graham, B. 2005. "Identifying Social Interactions Through Variance Contrasts." Unpublished.
- Graham, B. 2008. "Identifying Social Interactions Through Conditional Variance Restrictions." *Econometrica* 76: 643-660.
- Heckman, J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47: 153-161.
- Heckman, J. 2000. "Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective." *Quarterly Journal of Economics* 115: 45-97.
- Heckman, J. 2005. "The Scientific Model of Causality." *Sociological Methodology* 35: 1-98.
- Heckman, J. and R. Robb. 1985. "Alternative Methods for Evaluating the Impact of Interventions: An Overview." *Journal of Econometrics* 30: 239-267.
- Heckman, J. and R. Robb. 1986. "Alternative Methods for Solving the Problem of Selection Bias in Evaluation the Impact of Treatments on Outcomes." In *Drawing Inferences from Self-Selected Samples*, H. Wainer, ed. Berlin: Springer-Verlag.
- Hsiao, C. 1983. "Identification." In *Handbook of Econometrics*, Vol. 1. Z. Griliches and M. Intriligator, eds. Amsterdam: Springer Verlag.
- Hurwicz, L. 1966. "On the Structural Form of Interdependent Systems." In *Logic, Methodology and Philosophy of Science*, Proceeding of the 1960 International Congress, Studies in Logic and the Foundations of Mathematics, vol. 44. P. Suppes E. Nagel and A. Tarski, eds. Amsterdam: Elsevier.
- Ioannides Y. and J. Zabel. 2008. "Interactions, Neighborhood Selection, and Housing Demand." *Journal of Urban Economics*: 63: 229-52.
- Ising, E. 1925. "Beitrag zur Theorie des Ferromagnetismus." *Zeitschrift fur Physik*. 31: 253-258.
- Iyer, S. and M. Weeks. 2009. "Social Incentives, Ethnicity, and Fertility in Kenya," mimeo University of Cambridge.

Jackson, M. 2008. *Social and Economic Networks*. Princeton, NJ: Princeton University Press.

Kohler, H.-P., J. Behrman, and S. Watkins. 2001. "The Density of Social Networks and Fertility Decisions: Evidence from South Nyanza District, Kenya." *Demography* 38: 43-58.

Krauth, B. 2006. "Simulation-Based Estimation of Peer Effects." *Journal of Econometrics* 133: 243-271.

Lazarsfeld, P. and R. Merton. 1954. "Friendship as a Social Process: A Substantive and Methodological Approach." In *Freedom and Control in Modern Society*, M. Berger, ed. New York: van Nostrand.

Lee, L-f. 2007. "Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors and Fixed Effects." *Journal of Econometrics* 140: 333-374.

Lee, L-f., X. Liu, and X. Lin. 2009. "Specification and Estimation of Social Interaction Models with Network Structure, Contextual Factors, Correlation and Fixed Effects." *Econometrics Journal*, forthcoming.

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

Marschak, J. 1953. "Economic Measurements for Policy and Prediction." In *Studies in Econometric Method*, W. Hood and T. Koopmans, eds. New Haven: Yale University Press.

McPherson, M., L. Smith-Lovin, and J. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27: 415-444.

Miguel, E. and M. Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica*, 72: 159–217.

Moffitt, R. 2001. "Policy Interventions, Low-Level Equilibria, and Social Interactions." In *Social Dynamics*, S. Durlauf and H. P. Young, eds. Cambridge: MIT Press.

Munshi, K. 2003. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market." *Quarterly Journal of Economics* 118: 549-599.

Munshi, K. 2004. "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution." *Journal of Development Economics* 73: 185-213.

Nakajima, R. 2007. "Measuring Peer Effects on Youth Smoking Behaviour." *Review of Economic Studies* 74: 897-935.

Navarro, S. 2008. "Control Functions." In *The New Palgrave Dictionary of Economics*, second edition, S. Durlauf and L. Blume, eds. London: Palgrave MacMillan.

Nesheim, L. 2002. "Equilibrium Sorting of Heterogeneous Consumers across Locations: Theory and Empirical Implications". *CeMMAP Working Paper CWP08/02*.

Sirakaya, S. 2006. "Recidivism and Social Interactions." *Journal of the American Statistical Association* 101: 863-875.

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

Topa, G. 2001. "Social Interactions, Local Spillovers, and Unemployment." *Review of Economic Studies* 68: 261-295.