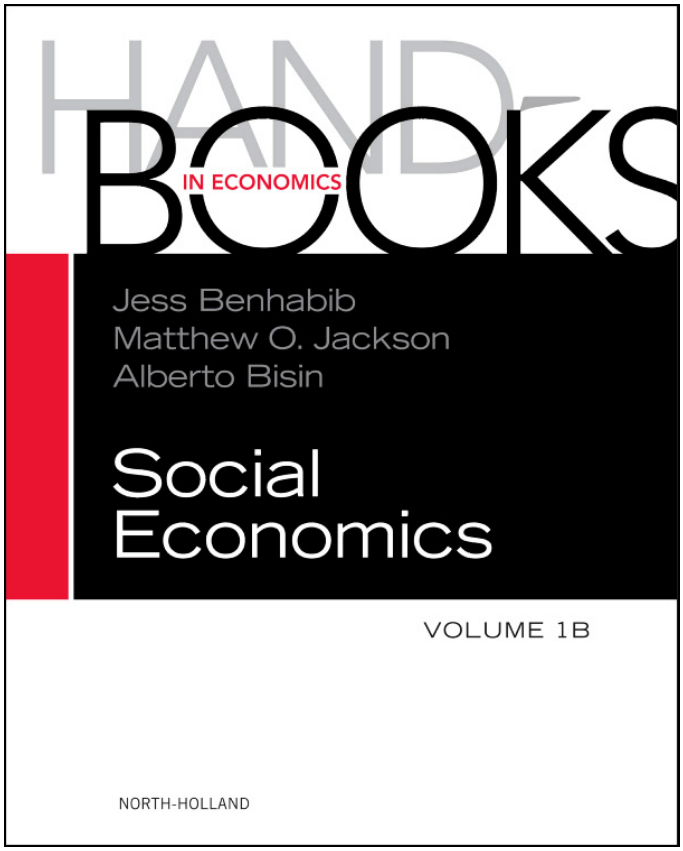

From: Lawrence E. Blume, William A. Brock, Steven N. Durlauf, and Yannis M. Ioannides, Identification of Social Interactions. In Jess Benhabib, Matthew O. Jackson and Alberto Bisin editors: Handbook of Social Economics, Vol. 1B,
The Netherlands: North-Holland, 2011, pp. 853-964.
ISBN: 978-0-444-53707-2

CHAPTER 18

Identification of Social Interactions[§]

Lawrence E. Blume,[*] William A. Brock,[†] Steven N. Durlauf,[**] and Yannis M. Ioannides[‡]

[*]Cornell University, IHS Vienna, Santa Fe Institute, lb19@cornell.edu
[†]University of Wisconsin wbrock@ssc.wisc.edu
[**]University of Wisconsin sdurlauf@ssc.wisc.edu
[‡]Tufts University yannis.ioannides@tufts.edu

Contents

[§] Prepared for the *Handbook of Social Economics*, J. Benhabib, A. Bisin, and M. Jackson, eds. We thank Andres Aradillas-Lopez, Jane Cooley, James Heckman, Charles Manski and Salvador Navarro for many conversations on the subject of this chapter. Financial support has been supplied to Blume by NSF grant CCF-0910940 and WWTF Grant "Die Evolution von Normen and Konventionen in der Wirtschaft", Brock 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, Xiangrong Yu and Yu Zhu have provided outstanding research assistance.
Corresponding Author: Steven N. Durlauf, Department of Economics, University of Wisconsin, Madison WI 53706.

ISSN 0169-7218, DOI: 10.1016/S0169-7218(11)01023-9

Abstract

While interest in social determinants of individual behavior has led to a rich theoretical literature and many efforts to measure these influences, a mature "social econometrics" has yet to emerge. This chapter provides a critical overview of the identification of social interactions. We consider linear and discrete choice models as well as social networks structures. We also consider experimental and quasi-experimental methods. In addition to describing the state of the identification literature, we indicate areas where additional research is especially needed and suggest some directions that appear to be especially promising.

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

Keywords

social interactions
social networks
identification

Like other tyrannies, the tyranny of the majority was at first, and is still vulgarly, held in dread, chiefly as operating through the acts of the public authorities. But reflecting persons perceived that when society is itself the tyrant-society, collectively over the separate individuals who compose it-its means of tyrannising are not restricted to the acts which it may do by the hands of its political functionary. Society can and does execute its own mandates: and if it issues wrong mandates instead or right, or mandates at all in things with which it ought not to meddle, it practices a social tyranny more formidable than many kinds of political oppression, since, though not usually upheld by such extreme penalties, it leaves fewer means of escape, penetrating more deeply into the details of life, and enslaving the soul itself. Protection, therefore, against the tyranny of the magistrate is not enough: there needs protection also against the tyranny of prevailing opinion and feeling; against the tendency of society to impose, by means other than civil penalties, its own ideas and rules of conduct on those who dissent from them; to fetter the development and, if possible prevent the formation, of any individuality not in harmony with its ways, and compel all characters to themselves upon the model of its own.

***John Stuart Mill,* On Liberty *(1859)*[1]**

1. INTRODUCTION

This chapter explores identification problems that arise in the study of social economics. We survey some of the existing empirical work, but do so in the context of different identification strategies. Our concern is understanding general conditions under which the finding of evidence of social interactions is possible and when it is not; we therefore do not focus on particular contexts. A valuable complement to our chapter is Epple and Romano (this volume) who provide an integration of theoretical, econometric, and empirical work on the specific question of peer effects in education.

While there now exists a rich literature which develops theoretical models of social interactions, as well as an enormous number of empirical papers that purport to find evidence for or against the presence of social interactions in particular contexts, the conditions under which social interactions are identified have yet to be comprehensively evaluated.[2] The main identification challenges facing an empiricist are now relatively well understood. These challenges come in three forms. One set of issues involves classical simultaneous equations problems. This set amounts to asking whether, for an equilibrium set of individual choices, one can differentiate between social interactions that derive

[1] On Liberty and other Writings, S. Collini ed., Cambridge: Cambridge University Press, p. 8.

[2] The empirical work is surveyed in Brock and Durlauf (2001b), Ioannides and Loury (2004), and Durlauf (2004). Theoretical models are surveyed elsewhere in this Handbook.

from direct interdependences between the choices and social interactions that derive from predetermined social factors. A second set of issues involves the possible presence of unobserved group-level characteristics. A third set of issues addresses the identification challenges that arise from the endogeneity of the groups that act as carriers of social interactions and the effects of self-selection. The research frontier in the study of social interactions thus involves efforts to achieve identification in the presence of these challenges. We will discuss a range of results on when point or partial identification holds.

These challenges all involve the canonical identification question for social interactions empirics: given an individual's membership in some group, can the analyst distinguish a role for the characteristics and behaviors of others in the group in influencing that individual's choices? We will thus generally speak of individuals as members of groups and describe social interactions in terms of those groups; the partial exception to this will occur when we consider networks which will characterize social structures across a population as a whole. In social interactions models, groups are typically defined in terms of exogenous categories such as ethnicity, gender or religion or endogenous categories such as residential neighborhoods, friendship networks, schools and firms. The former, of course, may not literally be exogenous, but rather the determination of whether an individual is a member of the category is treated as predetermined from the perspective of the behaviors under study. The distinction between exogenous and endogenous categories is of particular importance if one wants to use social interactions to understand a phenomenon such as inequality; without a model of how category memberships are determined a social economics theory of inequality will be incomplete.

Even for the case of exogenous categories, outstanding questions exist as to why particular categories act as the carriers of social influence whereas others do not. To be concrete, one could group individuals by eye color rather than race, yet there is no serious argument that the latter helps define peer interactions and the like. Issues of salience have received recent treatment in the economics of identity, e.g., Akerlof and Kranton (2000; 2002) and Austen-Smith and Fryer (2005). It seems evident that the salience of particular categories is history dependent. Loury (2002) makes a strong argument for this to be the case in terms of racial stigma in the United States.[3]

This chapter ignores two important dimensions of the econometrics of social interactions. First, we do not discuss issues of estimation, and refer the reader to papers such as Aradillas-Lopez (2010), Bajari, Hong, Krainer, and Nekipelov (forthcoming), Cooley (2008), de Paula (2009), Graham (2008), and Krauth (2006a), to see the progress that has been made in understanding the properties of various estimators. Second, we do

[3] The relationship between behaviors, group memberships and salience is, in our view, best understood in terms of time scales. In many cases, behavior choices are on a fast time scale relative to endogenous group membership. The time scale for membership is in turn fast relative to the process by which salience is determined. This perspective has the potential for unifying these three features of socioeconomic environments both with respect to theory and with respect to econometrics, but has yet to be pursued.

not directly address the estimation of treatment effects when social interactions are present. While we discuss the relationship between evidence of treatment effects from experimental interventions in group formations, we do not discuss the evaluation of evidence of intervention effects per se. Graham (2010) provides an extensive synthesis of work on such questions. Similarly, we do not discuss how social interactions affect the analysis of treatment effects; see Manski (2010) for recent work that addresses this question.

Section 2 of this chapter outlines a baseline model of decisionmaking in group contexts. Section 3 describes linear models of social interactions. Section 4 extends the analysis of linear models to social network and spatial interactions models. Section 5 describes discrete choice models of social interactions. Section 6 discusses identification in the context of experimental data. Section 7 proposes some dimensions along which we think new directions on the identification of social interactions ought to proceed, noting efforts that have already been made along these lines. Section 8 concludes.

We leave this introduction with a final observation. Loury's analysis of the historical specificity of stigma towards African Americans hints at a more general claim: empirical evidence in the social economics literature ought not to be limited to statistical studies. Within economics, a number of authors have produced persuasive evidence of social interactions via clever choices of environments to study. Examples include Costa and Kahn (2007) who show how a range of outcomes for prisoners in the U.S. Civil War are associated with social networks and Young and Burke (2001) who show how the terms of tenant contracts in Illinois cluster around a small set of simple fractional division rules even though these rules are unrelated to the sorts of contracts predicted by theory. Durlauf (2006) argues, more generally, that social psychology experiments, ethnographic studies, and historical analyses may well provide more persuasive evidence of social interactions than the existing body of formal statistical studies. We do not believe that the relative weakness of econometrically oriented studies must hold in principle. Rather, we wish to emphasize that econometric studies of social interactions are one part of a larger body of evidence that is relevant to their assessment and to their incorporation in policy evaluation. In fact, one reason why we regard the continuing study of identification problems for social interactions as so important is that statistically rigorous empirical work has, in our view, been the main source of empirical progress in the social sciences.

2. DECISION MAKING IN GROUP CONTEXTS

Our baseline model of social interactions studies the joint behavior of individuals who are members of a common group g. The population size of a group is denoted as n_g. Our objective is to probabilistically describe the individual choices of each i, ω_i. Choices are made from the elements of some set of possible behaviors Ω_{ig}. This set is both individual- and group-specific, though the econometric literature has typically not exploited the fact that different groups may offer different choices. This is an unexplored and interesting

possibility. For each i, ω_{-ig} denotes the choices of others in the group, which are one possible source of social interactions. From the perspective of econometric evaluation, it is useful to distinguish between five forms of influences on individual choices. These influences have different implications for how one models the choice problem. These forms are:

x_i	An R-vector of observable (to the modeler) individual-specific characteristics;
y_g	An S-vector of observable (to the modeler) group-specific characteristics;
$\mu_i^e(\omega_{-ig})$	A probability measure, unobservable (to the modeler), that describes the beliefs individual i possesses about behaviors of others in the group;[4]
ε_i	A vector of random individual-specific characteristics describing i, unobservable to the modeler; and
α_g	A vector of random group-specific characteristics, unobservable to the modeler.

The distinction between observable and unobservable determinants of individual choices corresponds to the standard difference between observable and unobservable heterogeneity in econometrics, or even more crudely, between the data ω_{ig}, x_i, y_g and the full range of factors affecting choices. Among the different sources of unobserved heterogeneity, $\mu_i^e(\omega_{-ig})$ functions very differently from ε_i and α_g since the logic of the choice problem determines the structure of $\mu_i^e(\omega_{-ig})$ in ways that do not apply to the other terms, which are shocks from the perspective of the modeler.

Individual choices ω_{ig} are characterized as representing the maximization of some payoff function V,

$$\omega_{ig} \in \operatorname{argmax}_{\lambda \in \Omega_{ig}} V(\lambda, x_i, y_g, \mu_i^e(\omega_{-ig}), \varepsilon_i, \alpha_g). \tag{1}$$

The decision problem facing an individual, a function of preferences (embodied in the specification of V); constraints (embodied in the specification of $\mathbf{\Omega}_{ig}$); and beliefs (embodied in the specification of $\mu_i^e(\omega_{-ig})$).[5] Thus, it is based on completely standard microeconomic reasoning. While the equilibria of these models can exhibit a range of interesting properties, such as multiple equilibria and bifurcations of the equilibrium properties of the environment around certain parameter values, these are properties of equilibria generated by this standard choice framework.[6]

As suggested above this choice model with social interactions is closed by the assumptions under which $\mu_i^e(\omega_{-ig})$ is determined. Without some structure on these

[4] Li and Lee (2009) consider the use of survey data to render beliefs observable; we discuss their work in section 5.vi. For purposes of the elucidation of the basic theory of choice in the presence of social interactions, we focus on the case where beliefs are latent variables.

[5] Throughout, probability measures are denoted by $\mu(\cdot)$.

[6] See surveys by Blume and Durlauf (2001), Brock and Durlauf (2001b) and Durlauf and Ioannides (2010) for overviews of these and other theoretical features of these models as well as the bibliographies of these papers for specific theoretical contributions.

beliefs, the model is consistent with any observed pattern of undominated choices. The standard assumption in the theoretical and econometric literatures, which we follow, closes the model by imposing an equilibrium condition: self-consistency between subjective beliefs $\mu_i^e(\omega_{-ig})$ and the objective conditional probabilities of the behaviors of others given i's information set F_i,

$$\mu_i^e(\omega_{-ig}) = \mu(\omega_{-ig}|F_i). \tag{2}$$

The requirement in (2) is usually called self-consistency in the social interactions literature and is nothing more than an equilibrium condition, from the perspective of empirical analysis. We assume that for each i, F_i consists of, for all $(x_j)_{j\in g}$, γ_g, ε_i and α_g. In other words, each agent knows his own characteristics x_i, as well as those of others in the group, the observed and unobserved group-level characteristics of his group (and of other groups), and his idiosyncratic error. Agents do not observe the ε_j's of others.

From the perspective of modeling individual behaviors, it is typically assumed that agents do not account for the effect of their choices on the decisions of others via expectations formation. The equilibrium in this model can be seen as a Bayes-Nash equilibrium of a simultaneous-move incomplete-information game. The individual decisions are described by

$$\omega_{ig} = \psi(x_i, \gamma_g, \mu(\omega_{-ig}|F_i), \varepsilon_i, \alpha_g). \tag{3}$$

Existence of an equilibrium for the group-wide vector of choices ω_g is equivalent to establishing that there exists a joint probability measure of these choices such that (3) is consistent with this joint probability measure. In applications in the literature, this is typically assured by a standard fixed point theorem, e.g., Brock and Durlauf (2001a), Cooley (2008). Notice that it is possible for γ_g and $\mu(\omega_{-ig}|F_i)$ to appear in equation (1) but not in equation (3). In this case, group behaviors and characteristics act as externalities but do not influence individual behaviors. This distinction is discussed in Cooper and John (1988). From the perspective of the empirical study of social interactions, equation (3) has been the main object of interest. Typically, (3) is assumed to exhibit a form of supermodularity in the sense that the redistribution of probability mass of $\mu(\omega_{-ig}|F_i)$ towards larger (in an element-by-element pairwise comparison sense) vectors of choices of others increases ω_{ig}. Milgrom and Roberts (1990) and Vives (1990) launched the now immense literature in economics on how supermodularity affects equilibrium outcomes for a wide range of environments; ideas from this literature often indirectly appear in the empirical social interactions literature, but with the exception of Aradillas-Lopez (2009), discussed in section 5.vi.d, this literature has been underutilized in the study of identification.

The distinction between γ_g and $\mu(\omega_{-ig}|F_i)$ is important in the social econometrics literature. Following Manski (1993), the former is known as a contextual effect whereas the latter (including the case of perfect foresight) is known as an endogenous effect. The importance of this distinction is that contextual interactions involve the interactions of predetermined (from the perspective of the model) attributes of one agent

affecting another whereas endogenous interactions allow for the possibility of simultaneity of interactions in individual outcomes.

To see how identification problems arise in a social interactions explanation of inequality, consider the stylized fact that the probability that a student graduates from high school is negatively associated with growing up in a poor neighborhood. Among the many possible explanations for this bivariate relationship are the following:

1. Heterogeneity in educational outcomes is determined by family-specific investment. Poor parents, following Becker and Tomes (1979) or Loury (1981), invest fewer resources in their children's education. If parental income is a sufficient statistic for parental investment, then the mechanism for lower graduation rates among poorer individuals is observable, constituting an element of x_i. The low graduation rate/poor neighborhood relationship is due to the interfamily correlation of low incomes that defines a poor neighborhood.
2. Effort choices by students depend on their assessments of the payoff to education. Poor neighborhoods contain distributions of role models that adversely affect educational choices. If a poor neighborhood tends to contain individuals whose incomes are relatively low compared to educational levels (as would occur via self-selection of lower incomes into poor neighborhoods), then the payoff to education may appear less attractive to high school students and thereby affect effort in high school as well as graduation decisions.[7] Relative to our candidate explanations, observed occupations, and educational levels of adults in a community are observable, and so are included in elements of γ_g. This is an example of how contextual effects can link poverty and low graduation rates.
3. High school graduation decisions are influenced by the choices of peers because of a direct desire to conform to the behaviors of others. Poorer neighborhoods have the feature that low values of the $\mu(\omega_{-ig}|F_i)$ are self-reinforcing, whereas high values of the $\mu(\omega_{-ig}|F_i)$ are self-reinforcing for more affluent neighborhoods. Thus endogenous social interactions can explain the relationship, although one has to be careful to explain why the peer interactions lead to lower graduation rates in poorer neighborhoods. We can offer three possible explanations: (i) The unique equilibrium could be characterized by a social multiplier that magnifies the consequences of income differences. (ii) In the spirit of Brock and Durlauf (2001a), there could be multiple equilibria in low-income neighborhoods but not in more affluent neighborhoods, because the poor may face lower marginal returns to education, which would magnify the influence of peer interactions relative to education returns in the equilibrium decision rule. (iii) There could be multiple equilibria for both high and low

[7] Streufert (2002) formalizes this type of idea and shows that the intuitive story just given is in fact oversimplified in the sense that the mapping from neighborhood levels of parental education/outcome relationships to student assessments of the returns to education may not lead to lower estimates of the returns in poorer neighborhoods but the story we describe is possible.

income neighborhoods, and an (unmodeled) selection mechanism could favor different equilibria in different neighborhoods in a manner correlated with income.

4. Parents transmit a host of skills to their children. Following Cunha and Heckman (2007) and Heckman (2007), suppose that poorer parents tend to have lower cognitive and noncognitive skills, which help to explain their lower socioeconomic status and are in turn transmitted to their children. This would imply that correlations among ε_i are the reason why poor neighborhoods have lower graduation rates. This is an example of correlated unobservables and is suggestive of the standard self-selection problem in econometrics.
5. Graduation decisions are affected by the quality of schools, where quality involves a host of factors ranging from the distribution of teacher ability to safety. Poorer neighborhoods have lower unmeasured school quality, then neighborhood poverty is a proxy for a low value of α_g, i.e., the graduation finding is caused by an unobserved group effect.

The bottom line is that each of the factors we have identified as determinants of individual outcomes can produce a relationship between individual outcomes and neighborhood characteristics, even when the mechanism is individual and not socially based. Of course, no economist would ever consider arguing that the fact that poor neighborhoods are associated with lower graduation rates speaks to any of these mechanisms per se. The identification question is whether these different explanations are distinguishable given the sorts of data that are available for analysis. It is this question that motivates the methods we describe.

We close this section with the observation that the behavioral model (3) cannot be nonparametrically identified without additional assumptions on structure. One reason for this is the possible existence of the unobserved group effects α_g which cannot be disentangled from elements of γ_g: Formally, there exist classes of models such that for any proposed function $\phi(\cdot)$ and associated choices of unobservables α_g, one can choose an alternative function ϕ' and alternative choice of unobservables α'_g such that all probability statements about the observables are identical. Brock and Durlauf (2007) show this for the binary choice model with social interactions which contains far more structure than (3), a model we will discuss in section 5 below.

Nonparametric identification may also fail even if one rules out unobserved group effects; Manski (1993) Proposition 3 gives various cases under which nonparametric identification fails for a version of the individual decision function equation (3). Specifically, Manski studies an environment in which the expected value of each person's choice is determined by[8]

$$E(\omega_{ig}|\gamma_g, x_i) = \phi(E(\omega_{ig}|\gamma_g), x_i). \tag{4}$$

[8] Appendix 1 contains an example of a model where this is a Bayes-Nash equilibrium condition.

Each individual is small relative to the population, producing the rational expectations equilibrium condition

$$\mathrm{E}(\omega_{ig}|\gamma_g) = \int \mathrm{E}(\omega_{ig}|\gamma_g, x) dF_{x|\gamma_g} \tag{5}$$

where $F_{x|\gamma_g}$ is the conditional distribution function of x_i in group g given γ_g. To say each individual is small is to say that knowledge of his own x_i does not affect the distribution function of individual characteristics within his group in a non-negligible way. For the joint model (4) and (5) one set of conditions under which nonparametric identification fails are 1) the solution to equation (5) is unique and 2) x_i is functionally dependent on γ_g. It is evident under these conditions that one cannot nonparametrically identify the separate effects of x_i and γ_g in determining ω_{ig} since differences in outcomes between two individuals with differences in x_i can always be attributed to the differences in their associated values of γ_g.

To make this example concrete, suppose that x_i is an individual's income and γ_g is the mean income of a residential neighborhood. Functional dependence would occur if neighborhoods were perfectly segregated by income, i.e., no neighborhood contained individuals with different incomes. For this case, it would impossible to distinguish the roles of individual and neighborhood incomes on outcomes since they would coincide. Less trivially, suppose that neighborhoods are fully segregated by income, which means that the empirical supports of incomes across neighborhoods never intersect. Suppose that individual income has no direct effects on outcomes whereas average neighborhood income has a monotonic effect on equilibrium outcomes. In this case, one could not distinguish an effect of neighborhood incomes on outcomes from the case where individual incomes directly affect outcomes, but do so in a step function fashion, where the jumps coincide with income levels that define the lower endpoints of the neighborhood income supports.

Manski also shows that identification will fail when x_i and γ_g are statistically independent. Nonidentification follows from statistical independence because $E(\omega_{ig}|\gamma_g)$ will not vary across groups, and so the effect cannot be distinguished from a constant term. An obvious example of this would occur if families were distributed across neighborhoods in such a way that each neighborhood had the same mean income in realization. Manski's result is in fact more general and is based on the observation that statistical independence implies that $E(\omega_{ig}|\gamma_g) = \int \phi((E(\omega_{ig}|\gamma_g), x) dF_x$, which by uniqueness means that $E(\omega_{ig}|\gamma_g)$ must be independent of γ_g.[9]

The implication of the cases where nonparametric identification fails is that identification of the models that fall into those cases will require various classes of assumptions that are not well motivated by economic theory, such as restrictions on

[9] The argument may be seen in Manski (1993, p. 539).

functional forms. One might argue that this means that social interactions empirics should focus on natural or quasi-natural experiments. We reject this position for several reasons. First, we strongly concur with general arguments made by Heckman (2000, 2005, 2008) on the nature of empirical economics. The types of empirical questions one cares about in studying social interactions, such as counterfactuals, are always properties of models and therefore require assumptions; good empirical work involves assessing factors such as robustness rather than rejecting assumptions per se. Second, the fact that an empirical claim is conditioned on an assumption that is not suggested by economic theory begs the question as to the degree of plausibility of the assumption. For example, in our earlier discussion of nonidentification of individual income versus residential neighborhood income effects, the model needed to explain outcomes exclusively as a function of individual incomes required that the mapping between the expected outcome and individual income follows a step function structure where the changes were coincident with income levels that defined the levels that separated neighborhoods. In our view, this is a relatively implausible assumption compared to a linear model. Again, these types of judgments are the stuff of science. Finally, we will argue that experimental and quasi-experimental analyses are also dependent on assumptions that are not justified by economic theory. Evidence of social interaction effects based on econometric models with nonexperimental data should be treated with a modesty, which reflects the degree of belief one is willing to place on identifying assumptions, but should not be dismissed altogether.

3. LINEAR MODELS OF SOCIAL INTERACTION

i. Basic structure

a. Description

Much of the empirical literature on social economics has involved variations of a general linear model, dubbed by Manski (1993) the linear-in-means model

$$\omega_{ig} = k + cx_i + d\gamma_g + Jm^e_{ig} + \varepsilon_i, \tag{6}$$

where m^e_{ig} denotes the average behavior in the group, i.e.,

$$m^e_{ig} = \frac{1}{n_g}\sum_{j\in g} \mathrm{E}(\omega_j|F_i). \tag{7}$$

Following our definitions of the variables, note that k and J are scalars whereas c and d are R- and S-vectors, respectively.[10] Claims about social interactions are, from the econometric perspective, equivalent to statements about the values of d and J. The statement that social interactions matter is equivalent to the statement that at least

[10] Throughout, coefficient vectors such as c are row vectors whereas variable vectors such as x_i are column vectors.

some element of the union of the parameters in d and the scalar J are nonzero. The statement that contextual social interactions are present means that at least one element of d is nonzero. The statement that endogenous social interactions matter means that J is nonzero. In Manski's original formulation, $\gamma_g = \bar{x}_g$, where $\bar{x}_g = \frac{1}{n_g}\sum_{j\in g} x_j$ denotes the average across individuals i of individual characteristics x_i within a given group g, which explains the model's name. Regardless of whether they are equal, we assume that both γ_g and $\bar{x}_g$ are observable to individuals, and discuss how to relax this below.

The linear in means model is typically invoked without any explicit attention to individual decision problems and associated equilibria as described in section 2. Appendix 1 provides an explicit derivation of the model from individual decision problems and shows how (6) describes the unique decision rules in a Bayes-Nash equilibrium under a particular functional form assumption on individual preferences.

We initially study the model under two assumptions on the errors. First we assume that the expected value of ε_i is 0, conditional on the information set $(x_i, \bar{x}_g, \gamma_g, i \in g)$,[11]

$$\text{for each } g \text{ and } i \in g \quad E(\varepsilon_i | x_i, \bar{x}_g, \gamma_g, i \in g) = 0. \tag{8}$$

Second we assume that

$$\begin{gathered}\text{for each } i, j, g, h \text{ such that } i \neq j \text{ or } g \neq h \\ \operatorname{cov}(\varepsilon_i \varepsilon_j | x_i, \bar{x}_g, \gamma_g, i \in g, x_j, \bar{x}_h, \gamma_h, j \in h) = 0.\end{gathered} \tag{9}$$

Equation (9) eliminates conditional covariation between the errors. The inclusion of the group memberships, e.g., $i \in g$, rules out some relationship between the identity of the group and model errors, thereby allowing us to treat groups as exchangeable.

From equations (6) and (7), and assuming that each individual is small enough relative to the group that the effect of his knowledge of his own ε_i on m^e_{ig} can be ignored, equilibrium implies that each actor's expected average behavior will be equal to a common value. This common value is derived in appendix 1 and is described by

$$m^e_{ig} = m_g \equiv \frac{k + c\bar{x}_g + d\gamma_g}{1 - J}. \tag{10}$$

This equation says that the individuals' expectations of average behavior in the group equal the average behavior of the group, and this in turn depends linearly on the average of the individual determinants of behavior, $\bar{x}_g$, and the contextual interactions that the group members experience in common, γ_g. The condition $J < 1$, which is required for equation (10) to make sense, is guaranteed to hold in the game-model of appendix 1. There, J maps the marginal rate of substitution between private return and social conformity, a non-negative real number, into the interval [0, 1).

[11] The conditioning argument $i \in g$ means that one is conditioning on the fact that i is a member of group g.

b. Reduced form

Substitution of (10) into (6) eliminates m_g and so provides a reduced form version of the linear in means model in that the individual outcomes are determined entirely by observables and the individual-specific error:

$$\omega_{ig} = \frac{k}{1-J} + cx_i + \frac{J}{1-J}c\bar{x}_g + \frac{d}{1-J}\gamma_g + \varepsilon_i. \tag{11}$$

Much of the empirical literature has ignored the distinction between endogenous and contextual interactions, and has focused on this reduced form, i.e., focused on the regression

$$\omega_{ig} = \pi_0 + \pi_1 x_i + \pi_2 \gamma_g + \varepsilon_i, \tag{12}$$

where the parameters π_0, π_1, π_2 are taken as the objects of interest in the empirical exercise. A comparison of (12) with (11) indicates how findings in the empirical literature that end with the reporting of π_0, π_1, π_2 inadequately address the task of fully characterizing the social interactions that are present in the data. For example, from the perspective of (12), the presence of social interactions is equivalent to $\pi_2 \neq 0$, whereas from the perspective of (6) this is neither necessary nor sufficient for social interactions to be present since $J = 0$ is neither necessary nor sufficient for $\pi_2 = 0$. To be clear, this observation does not mean that estimates of (12) are uninformative, rather that these estimates should be mapped to structural parameters in the sense of (6) when identification holds, and that if identification does not hold, then the informational limits of (12) in terms of distinguishing types of social interactions should be made explicit.

Equation (12) is nonetheless the source of much of the current econometric evidence on social interactions. Datcher (1982) should be regarded as a seminal contribution to the social economics literature since it appears to be the first empirical study to propose (11) as an estimating equation. Her empirical specification has generally remained the empirical standard despite the importance that Manski (1993) subsequently attached to the contextual versus endogenous distinction in econometric work. An exception is Gaviria and Raphael (2001), although they do this by arguing that contextual interactions are not relevant for their context, teen behaviors.

The reduced form version of the linear in means model illustrates some features of the structure that are of interest. First, the linear in means model limits the effects of reallocations of individuals across groups. To see this, suppose one thinks of each choice in the population as $\omega_{ig} = \phi(x_i, \gamma_g) + \varepsilon_i$. Suppose that γ_g is a scalar and that $\gamma_g = \bar{x}_g$. This means that $\omega_{ig} = \phi(x_i, x_{-ig}) + \varepsilon_i$, where x_{-ig} denotes the vector of individual characteristics other than x_i among group g members with typical element x_{-ijg}. Finally, assume all groups are of equal size. Under the linear functional form (11), for all j, $\partial^2\phi/\partial x_i x_{-ijg} = 0$. This is the condition under which all allocations of individuals across groups produce the same expected population-wide average outcome for $\sum_{j\in g}\omega_{jg}$. This was first

recognized in Becker's (1973) analysis of efficiency in the marriage market, in which groups are of size 2 and naturally extends to groups of any size. (See Durlauf and Seshadri (2003).) It is the case, extending an example of the type in Durlauf and Seshadri (2003), that if groups are of different sizes, the reallocation of individuals across them can affect average outcomes. This nonetheless does not diminish the qualitative point that the fact that all cross partial derivatives equal 0 in the reduced form of the linear in means model severely restricts the effects of reallocations of group memberships.

Second, if policies are available to influence x_i and/or γ_g, then these interactions can be identified even if the structural parameters are not identified. This observation is elaborated in Manski (2010) who emphasizes the distinction between structural model identification and potential outcomes identification, which lies at the heart of treatment effect analysis. Another way to think about this distinction is that for many policy contexts, the structural model is of no intrinsic interest. Brock, Durlauf, and West (2003) argue that this is the case for a range of macroeconomic contexts. However, this type of identification will not allow the policy analyst to address issues where the endogenous social interactions are themselves of fundamental policy relevance, as may be the case if these interactions alter the distribution of individuals across groups.

ii. Instrumental variables and the reflection problem

We first consider the estimates of the regression coefficients for (6) under the expectations formation restriction (10). It is obvious that if $\bar{\omega}_g$ is projected against the union of elements of $\bar{x}_g$ and γ_g, this produces the population mean m_g. Hence, we can proceed as if m_g is observable. Put differently, our identification arguments rely on the analogy principle which means that one works with population moments to construct identification arguments.[12] Since γ_g appears in (10), it will not facilitate identification. As we shall see, identification via instrumental variables is determined by the informational content of $\bar{x}_g$ relative to γ_g.

As first recognized by Manski (1993), identification can fail for the linear in means model when one focuses on the mapping from reduced form regression parameters to the structural parameters. This may be most easily seen under Manski's original assumption that $\gamma_g = \bar{x}_g$. This means that every contextual effect is the average of a corresponding individual characteristic. In this case, equation (10) reduces to

$$m_g = \frac{k + (c + d)\gamma_g}{1 - J}. \tag{13}$$

This means that the regressor m_g in equation (6) is linearly dependent on the other regressors, i.e., the constant and γ_g. This linear dependence means that identification fails: the comovements of m_g and γ_g are such that one cannot disentangle their respective influences on individuals. Manski (1993) named this failure the reflection problem.

[12] Goldberger (1991, p. 117) gives a concise description.

Metaphorically, if one observes that ω_{ig} is correlated with the expected average behavior in a neighborhood, (13) indicates it may be possible that this correlation is due to the fact that m_g may simply reflect the role of γ_g in influencing individuals.

Under what conditions is this model identified? A necessary condition is that Manski's assumption that $\gamma_g = \bar{x}_g$ is relaxed. This will allow for the possibility m_g is not linearly dependent on the constant and γ_g. The reason for this is the presence of the term $c\bar{x}_g/(1-J)$ in equation (10). This term can break the reflection problem. This will happen if the $c\bar{x}_g/(1-J)$ term is not linearly dependent on a constant and γ_g. When this is so, m_g cannot be linearly dependent on the other regressors in equation (10). This immediately leads to the argument in Brock and Durlauf (2001b) that a necessary condition for identification in the linear in means model, is that there exists at least one element of x_i whose group level average is not an element of γ_g, while Durlauf and Tanaka (2008) provide a sufficient set of conditions. Necessity and sufficiency can be linked as follows. Let proj$(a|b, c)$ denote the linear projection of the scalar random variable a onto the elements of the random vectors b and c.[13] Consider the projections proj$(\omega_g|1, \gamma_g, \bar{x}_g)$ and proj$(\omega_g|1, \gamma_g)$, where 1 is simply a random variable with mean 1 and variance 0, corresponding to the constant term in (6). The first projection provides an optimal linear forecast (in the variance minimizing sense) of the group average choice, $\bar{\omega}_g = (1/n_g) \sum_{i \in g} \omega_{ig}$, conditioning on the random variables defined by 1 and the elements of γ_g and $\bar{x}_g$, whereas the second projection provides the optimal linear forecast when only 1 and the elements of γ_g are used. The difference between the two projections thus measures the additional contribution to predicting $\bar{\omega}_g$ beyond what can be achieved using $\bar{x}_g$ in addition to 1 and γ_g. When this marginal contribution is nonzero, then it is possible to estimate equation (10) using instrumental variables for $\bar{\omega}_g$ or equivalently estimate (6) when (10) is imposed by instrumenting m_g.[14] Formally,

Theorem 1. Identification in the linear in means model. *Identification of the parameters k, c, J and d requires that* proj$(\bar{\omega}_g|1, \gamma_g, \bar{x}_g) -$ proj$(\bar{\omega}_g|1, \gamma_g) \neq 0$.

The intuition for the theorem is simple; identification requires that one can project $\bar{\omega}_g$ (equivalently) onto a space of variables such that the projection is not collinear with the other regressors in the model. As such, the theorem verifies that identification in the linear in means model is a species of identification of a linear simultaneous equations system, as argued above.[15]

Theorem 1 was derived under the assumption that $\bar{x}_g$ and γ_g are known to the individual decisionmakers at the time that their choices are made. This assumption is a

[13] Formally, this is the projection of a onto the Hilbert space generated around the elements of b and c where the inner product between any two elements is the expected value of their product so that the metric measuring the length of an element is the square root of the inner product of an element with itself.

[14] Recall that in equilibrium, proj$(\bar{\omega}_g|1, \gamma_g, \bar{x}_g) =$ proj$(m_g|1, \gamma_g, \bar{x}_g)$ and proj$(\bar{\omega}_g|1, \gamma_g) =$ proj$(m_g|1, \gamma_g)$.

[15] The conditions of the theorem do not preclude a functional dependence of x_i on γ_g, which, combined with the uniqueness of m_g, means that the nonparametric analog to the model is not identified, following Manski (1993, Proposition 3). This observation builds on discussion in Manski (1993, p. 539).

strong one and further may appear to be inconsistent with our assumption that $\bar{\omega}_g$ is unobservable to them. This latter concern is not tenable: in a context such as residential neighborhoods, it is possible for a contextual effect such as average income to be observable whereas the school effort levels of children in the neighborhood are not. However, it is important to understand the implications of relaxing informational assumptions on identification. This is the contribution of Graham and Hahn (2005). The models they study can be subsumed as variants of a modified version of equation (6):

$$\omega_{ig} = k + cx_i + d\mathrm{E}(\gamma_g|F) + Jm_g + \varepsilon_i \tag{14}$$

where individuals are assumed to possess a common information set F. As such, it is clear that the conditions for identification in theorem 1 are easily generalized. One simply needs a set of additional instruments q_g such that the elements of q_g can jointly instrument $\mathrm{E}(\gamma_g)$ and m_g. As they observe, the variables q_g constitute exclusion restrictions and so require prior information on the part of the analyst. For their context, γ_g is a strict subset of $\bar{x}_g$, so it is difficult to justify the observability of those elements of $\bar{x}_g$ that do not appear in γ_g when the others are by assumption not observable. In our view, the appropriate route to uncovering valid instruments q_g, under the Graham and Hahn information assumptions, most likely requires the development of an auxiliary model of x_i and hence $\bar{x}_g$. In other words, Graham and Hahn's concerns reflect the incompleteness of (14) in the sense that the individual characteristics are not themselves modeled. Hence, we interpret their argument as one that calls for the embedding of outcomes such as (14) in a richer simultaneous equation system, possibly one including dynamics, which describes how individual characteristics are determined. We fully agree with Graham and Hahn that in isolation, finding valid instruments for (14) is difficult, but would argue that this difficulty reflects the limitations of studying ω_{ig} in isolation rather than as one of a set of equilibrium outcomes.

We now evaluate the reflection problem for some econometric models that differ from (6) in various ways that are common in empirical work. Once one considers econometric structures outside the linear cross-section framework, the reflection problem may not arise, even if there is a one-to-one correspondence between individual and contextual interactions. We consider three alternative structures.

a. Partial linear in means models

The linear structure in (6) is typically only theoretically justified under strong functional form assumptions for utility, as shown in appendix 1, which leads to the question of whether relaxation of the linearity assumption affects identification. One such relaxation is studied in Brock and Durlauf (2001b) and involves a particular nonlinear generalization of (6) under rational expectations

$$\omega_{ig} = k + cx_i + d\gamma_g + J\mu(m_g) + \varepsilon_i. \tag{15}$$

This type of structure is known as a partial linear model. Brock and Durlauf establish that the parameters of this model are identified for those elements of the space of twice differentiable functions, for known $\mu(m_g)$, so long as $\partial^2\mu(m_g)/\partial m_g^2 \neq 0$, outside of nongeneric cases. The intuition is straightforward; the reflection problem requires linear dependence between group outcomes and certain group-level aggregates, which is ruled out by the nonlinearity in (15). Note that there does not exist any identification results, as far as we know, if the functional form for $\mu(m_g)$ is unknown, so in this sense the identification of (15) does not exploit results from the semiparametric literature on partial linear models.[16]

The finding that partial linear variants of (6) do not suffer from the reflection problem is not a surprise from the perspective of the simultaneous equations literature. McManus (1992), in what appears to be an underappreciated paper, illustrates how for a broad class of parametric nonlinear simultaneous equations models, subsets of nonidentified models are nongeneric. For example, McManus (1992, p. 8) shows in his pedagogical example that "...First the set of δ values which correspond to identified (non identified) models forms an open and dense (nowhere dense) subset of the real line..." He develops a general argument which formalizes this basic idea. Brock and Durlauf (2001b, p. 3371) adapt McManus's argument to show that "...the local nonidentification of the linear-in-means model can be perturbed away by a C^2-small change." See Brock and Durlauf (2001b, Chap. 54) for the details of this extension to social interaction models.

This example of generic identifiability of a nonlinear in means model illustrates the importance of treating the quadratic utility function in appendix 1 as the true preference specification as opposed to regarding it as some sort of second-order approximation. As a second order approximation, the preference specification leads to erroneous conclusions about identification. Interestingly, a corresponding set of findings have been developed by Ekeland, Heckman, and Nesheim (2004) in the context of hedonic models. These authors show that previous claims of lack of identification in hedonic models are special to linear specifications that derive from a quadratic utility specification. While the arguments as to why identification holds for preferences outside the quadratic case, the hedonic context provides another demonstration of how the use of quadratic approximations can lead to misleading conclusions on identifiability for economically important environments.

b. Dynamic linear models

Similarly, dynamic analogs of the linear in means model may not exhibit the reflection problem. Brock and Durlauf (2001b) illustrate this with the dynamic social interactions model

$$\omega_{igt} = k + cx_{it} + dy_{gt} + \beta m_{gt-1} + \varepsilon_{it},$$

[16] See Tamer (2008) for a survey.

where for all s, $t \neq 0$,

$$\operatorname{cov}(\varepsilon_{it}, \varepsilon_{it-s}) = 0. \tag{16}$$

This model avoids linear dependence between the contextual and endogenous variables since

$$m_{gt} = \frac{k + c\bar{x}_{gt} + d\gamma_{gt}}{1 - \beta L} \tag{17}$$

where L is a lag operator. Equation (17) implies that m_{gt} depends on the entire history of $\bar{x}_{gt}$ and γ_{gt}. This model is essentially backwards looking and is driven by the idea that current behaviors are directly affected by past beliefs. A more natural approach, of course, is to consider how beliefs about the future affect current behaviors. An example of a model in this class is

$$\omega_{igt} = k + cx_{it} + d\gamma_{gt} + \beta m_{gt+1} + \varepsilon_{it} \tag{18}$$

where (16) is again assumed. This model is equivalent to the workhorse geometric discount model in rational expectations (Hansen and Sargent, 1980).

The equilibrium average choice level for a group equals, following Hansen and Sargent,[17]

$$m_{gt} = \frac{k}{1 - \beta} + \sum_{s=0}^{\infty} \beta^s \mathrm{E}_t(c\bar{x}_{gt+s} + d\gamma_{gt+s}). \tag{19}$$

It is immediate from (19) that the regressors in (18) are linearly independent so long as $\bar{x}_{gt}$ and γ_{gt} are not both random walks. Identification of this class of dynamic models was originally studied in Wallis (1980) and has recently been explored in Binder and Pesaran (2001).

c. Hierarchical models

In fields such as sociology, social interactions are typically explored using hierarchical models, i.e., models in which contextual interactions alter the coefficients that link individual characteristics to outcomes. See Bryk and Raudenbush (2001) for a full description of the method. The reason for this appears to be a different conceptualization of the meaning of social interactions in economics in comparison to other social sciences. Hierarchical models appear, in our reading, to be motivated by a view of social groups as defining ecologies in which decisions are made and matter because different social backgrounds induce different mappings from the individual determinants of these behaviors and choices, cf. Raudenbush and Sampson (1999). Economics, in contrast, regards the elements that comprise endogenous and contextual social interactions as directly affecting the preferences, constraints, and beliefs of agents and so treats them as

[17] In this formulation we restrict ourselves to fundamental solutions of the expected average choice level. The possibility of a nonfundamental solution, i.e., bubbles, is not germane to the discussion.

additional determinants to individual specific characteristics, x_i. That said, there do not exist formal arguments for favoring one approach versus another at an abstract level. At the same time, the additivity assumption in both approaches is ad hoc from the perspective of economic theory, even if the assumption is ubiquitous in empirical practice.

For hierarchical models, there has been no attention to the reflection problem. The only exception of which we are aware is Blume and Durlauf (2005). Here we modify the Blume and Durlauf analysis and consider a formulation that closely follows the conceptual logic of hierarchical models in that social interactions are entirely subsumed in the interactions on parameters. Formally, this means that individual outcomes obey

$$\omega_{ig} = k_g + c_g x_i + \varepsilon_i \tag{20}$$

with individual- and group-specific components obeying

$$k_g = k + d\gamma_g + Jm_g \tag{21}$$

and

$$c_g = c + \gamma'_g \Psi + m_g \psi \tag{22}$$

respectively. In (22), Ψ is a matrix and ψ is a vector. We omit any random terms in (21) and (22) for simplicity, although hierarchical models typically include them. This formulation assumes that the endogenous effect directly affects the individual level coefficients and so differs from the Blume and Durlauf example. Imposing rational expectations, the hierarchical model described by (20)-(22) is equivalent to the linear model

$$\omega_{ig} = k + cx_i + d\gamma_g + Jm_g + \gamma'_g \Psi x_i + m_g \psi x_i + \varepsilon_i. \tag{23}$$

Hence, the difference between the linear model used in economics and the hierarchical structure is the addition of the terms $\gamma'_g \Psi x_i$ and $m_g \psi x_i$ by the hierarchical model to equation (6). Thus, the hierarchical model does nothing deeper than add the cross products of variables in (6) to allow for nonlinearity. As such, the approach is far behind the econometrics literature on semiparametric methods, which allows for much deeper forms of nonlinearity. On the other hand, the use of cross products of variables is still common in empirical economics.

Can this model exhibit the reflection problem? The self-consistent solution to (23) is

$$m_g = \frac{k + c\bar{x}_g + d\gamma_g + \gamma'_g \Psi \bar{x}_g}{1 - J - \psi \bar{x}_g}. \tag{24}$$

Recall that the reflection problem necessarily emerged in (6) when $\gamma_g = \bar{x}_g$. If we impose this condition in the hierarchical model, (24) becomes

$$m_g = \frac{k + (c + d)\gamma_g + \gamma'_g \Psi \gamma_g}{1 - J - \psi \gamma_g}. \tag{25}$$

Equation (25) makes clear that the relationship between m_g and the other regressors is nonlinear. Further, the presence of $\gamma'_g \Psi \gamma_g$ in the numerator and $-\psi\gamma_g$ in the denominator ensures that linear dependence will not hold, except for hairline cases, so long as there is sufficient variation in x_i and γ_g.

Hierarchical models thus exhibit different identification properties from linear in means models because their structure renders the endogenous effect m_g a nonlinear function of the contextual interactions γ_g (and also a nonlinear function of $\bar{x}_g$ if this variable is distinct from γ_g). The reflection problem can thus be overcome without prior information about the relationship between $\bar{x}_g$ and γ_g. However, this does not mean that users of hierarchical models of social interactions can ignore the possibility of endogenous social interactions and only focus on contextual effects. The nonlinear relationship between m_g and γ_g means that the failure to account for endogenous social interactions in hierarchical models will lead to inconsistent estimates of the contextual effect parameters. Further, hierarchical models cannot be used to evaluate the interactions of changes in different variables, or the interactions on individual outcomes of altering group memberships, e.g., by changing school district boundaries.[18] These types of policy interventions will depend on the value of all the social interactions parameters and the attendant nonlinearity described by equation (25). Hierarchical models thus contrast with the linear in means example given in Manski (2010) where policy evaluation does not require knowledge of all parameters.

Do these cases imply that the reflection problem is a special one, i.e., that it only arises for cross-sectional data and linear models which together render instrumental variable requirements difficult? We believe that this would be an incorrect assessment of the importance of Manski's result. While alternative specifications may imply that identification holds per se, it may also be the case that m_g is highly correlated with combinations of the other determinants of individual outcomes. Even if identification holds, parameter estimation may be highly imprecise, for the exact same reasons that the reflection problem leads to nonidentifiability in some contexts. Therefore, we regard the reflection problem as a fundamental difficulty in estimating social interactions.

iii. Variance-based approaches

As noted above, a second route to identification of the linear in means model may be derived from the covariance structure model errors. This approach is discussed in classic treatments of identification such as Fisher (1966) and relies on strong prior information on the covariance structure of a given model's errors. In general, this approach to identification became unpopular in economics because modern econometrics has emphasized the relaxation of assumptions on error structures, as manifested in the work on

[18] The argument we make with respect to assessing changes in group memberships is an example of the discussion found in section 3.i.b concerning the special implications of the linear in means model for the aggregate effects of the reallocation of individuals across groups.

heteroskedastic and autocorrelation consistent covariance matrix estimation initiated by White (1980).[19] This emphasis on econometric analysis under weak assumptions on errors is properly regarded as a major breakthrough since in many socioeconomic contexts, assumptions such as homoskedasticity have no theoretical justification. To the extent that theory does constrain the stochastic processes for model errors, modern econometrics has focused on incorporating this dependence into the empirical analysis. Heckman (2001) gives an overview of this perspective for microeconomics, which is of course the locus of social interactions. It is therefore unsurprising that most work on empirical social economics has avoided exploiting covariance restrictions as a source of identification.

That said, one can imagine contexts where strong assumptions are relatively appealing. One example is randomized experiments in which exchangeability of the individuals can be invoked and so assumptions of independence and identical errors can be justified by appeal to deFinetti's Theorem, which in essence says that the probability measure for an infinite exchangeable sequence of random variables can be written as a mixture of joint probability measures for i.i.d. random variables; since one of the components of the mixture is realized in the data, we can treat the data as if it were i.i.d.[20] This is the sort of context that motivated a strategy developed in Graham (2008) to uncover social interactions via covariance restrictions for Project Star, an experiment designed to study the consequences of classroom size. In Project Star students were randomly distributed across classes of different sizes; teachers were randomly assigned as well.

To see how this approach, which Graham refers to as the method of variance contrasts, works, we employ a simplified version of his model, which assumes that individual outcomes are affected by the realized mean outcomes in classrooms,[21]

$$\omega_{ig} = J\bar{\omega}_g + \varepsilon_i = Jm_g + \varepsilon_i + J\bar{\varepsilon}_g. \tag{26}$$

Individual and contextual interactions are thus assumed away, which renders the instrumental variable strategies we have described for identification impossible. Graham further assumes that the individual errors obey

$$\text{var}\,(\varepsilon_i | i \in G) = \sigma_\varepsilon^2 \tag{27}$$

and

$$\text{for } i \neq j, \;\; \text{cov}\,(\varepsilon_i \varepsilon_j | i, j \in g) = 0. \tag{28}$$

Graham's insight, which builds on earlier work by Glaeser, Sacerdote, and Scheinkman (1996) (which will be discussed in the context of discrete choice models) is that the presence of J affects the variance of ω_{ig} and may be used for identification. For this

[19] See West (2008) for an overview.
[20] See Bernardo and Smith (1994, section 4.2) for discussion.
[21] Graham allows for unobserved group level effects, which we consider below.

model, one can think of the outcomes as generated by a reduced form in which the errors fully determine the outcomes, in other words, all information about the model parameters is embedded in the variance covariance matrices of the various ω_g's. Graham shows that for the model (26), under assumptions (27) and (28):

$$\text{var}\,(\omega_g) = \left(I_{n_g} - \frac{J}{n_g}\iota_{n_g}\right)^{-2}\sigma_\varepsilon^2, \tag{29}$$

where I_{n_g} is an $n_g \times n_g$ identity matrix and ι_{n_g} is a $n_g \times n_g$ matrix of 1's. Equation (29) implies that if there are two groups with different sizes, one can use the differences in the intergroup outcome variances to identify J. Following Graham, this result follows intuitively from the fact that for larger groups the variance in $\bar{\omega}_g$ is smaller. We should note that the assumption expressed by (28) is stronger than what can be justified by exchangeability of the individual errors per se. Durlauf and Tanaka (2008) explicitly show that Graham's results follow if one starts with exchangeability of the individual errors and further assumes that error variances are independent of classroom size.

iv. Unobserved group effects

As suggested in the introduction, one of the major limits to identification of social interactions is the presence of unobserved group-level heterogeneity. To introduce this factor, we modify (6) to

$$\omega_{ig} = k + cx_i + d\gamma_g + Jm_g + \alpha_g + \varepsilon_i, \tag{30}$$

where rational expectations is imposed as in (10). The associated reduced form for (30) is

$$\omega_{ig} = \frac{k}{1-J} + cx_i + \frac{J}{1-J}c\bar{x}_g + \frac{d}{1-J}\gamma_g + \frac{1}{1-J}\alpha_g + \varepsilon_i. \tag{31}$$

It is evident from (31) that correlation of α_g with the regressors in the equation can lead to identification problems. It is hard to see how one can rule such correlations out. For example, correlation with x_i naturally arises from self-selection and correlation with γ_g naturally arises from imprecise or incomplete measurement of group contextual effects.

a. Instrumental variables

One approach to dealing with unobserved group level heterogeneity in (30) or (31) is the use of instrumental variables. This approach is generally difficult to justify in addressing unobserved group characteristics for both the linear in means and other models. The reason for the difficulty is that α_g is itself undertheorized, in other words, this term captures aspects of a group that affect outcomes which the model does not explicitly describe. Beyond this, valid instrumental variables require the property that they have been excluded from (30) as either individual or contextual determinants of outcomes. It is hard to see how, in typical socioeconomic contexts, such instruments may be found, since the

instruments must be known on *a priori* grounds to be uncorrelated with both the undertheorized α_g and ε_i. Social interactions models are typically what Brock and Durlauf (2001c) have termed openended, which means that their theoretical structure does not naturally identify variables to exclude from equations such as (30). In other words, social interactions theories are openended because the presence of a given type of social interaction does not logically preclude the empirical relevance of other theories; the econometric analog of this is that social economics models do not provide a logical basis for choosing instruments. This is quite different from rational expectations models, for example, whose logic often allows one to express linear combinations of variables as forecast errors, which must logically be orthogonal to an agent's information set; in macroeconomics a key example of this is the Euler equation in a stochastic optimization model.

We suspect that theory openendness for a given social interactions model, combined with an undertheorized α_g, have (at least implicitly) rendered instrumental variables strategies relatively uncommon in social interaction contexts, for both the linear in means and other models. One exception is Cooley (2008) who makes a careful substantive argument to justify her use of instruments. Cooley's objective is to estimate the effects of peer educational outcomes on individual students in a classroom; for her context teacher quality is an unobserved group effect. Cooley studies reading achievement in North Carolina and proposes an identification strategy that exploits an education reform enacted in that state that went into effect in the 2000-2001 academic year. The reform mandated that students who scored below a certain level on end of year reading tests are to be held back for the next year. For a rich panel data set on classrooms, Cooley proposes as an instrument the percentage of students who are held accountable (i.e., the percentage of students who scored below the cutoff in the prior year interacted with a dummy for being a 5th grader in 2001 or later). Cooley argues that these "in danger of failing" students face incentives to work harder under the educational reform. Classes with a higher percentage of low-achievers would see a larger shift in peer achievement after accountability. She thus contrasts comparable (from the perspective of student composition) classrooms before and after the educational reform to estimate a peer effect, with 4th grade classrooms acting as a control group.[22]

An obvious objection to this strategy is that for a given set of classroom student characteristics, unobserved teacher quality may have shifted after the reform, as teachers adjusted their behavior in response to the reform. Cooley addresses this objection by appealing to the history of educational reform in North Carolina and argues that a post 2000 shift in teacher behavior is relatively implausible. The 1996 reforms included school and teacher bonuses for growth in scores and for the percentage of students who met or exceeded the reading level that later was used to determine whether a grade should be repeated. Thus it is plausible to claim that adjustments in teacher behavior most probably date from then. Of course, Cooley does not "prove" that her instrument is valid. Our

[22] We thank Jane Cooley for helping with this description.

view is that that she has made a case that the instrument is a plausible one so that her findings should update researchers' beliefs about classroom peer effects. We believe this careful type of argumentation can be replicated in other contexts.

Some uses of instrumental variables fall under the rubric of quasi-natural experiments. A recent example is Cipollone and Rosolia (2007) which we describe in some detail as it illustrates the strengths and weaknesses of quasi-experimental data as a source for evidence on social interactions. Their analysis examines the effects of changes in male high school graduation rates on female high school graduation rates using a change in Italy's compulsory military service laws which exempted male students in schools located in areas damaged by a 1980 earthquake. Cipollone and Rosolia compare two groups of schools. The first group of schools are located in towns that experienced relatively little earthquake damage (based on official assessments) and yet were included in the draft exemption. The second group of schools were located in towns that were near the towns whose schools comprise the first group; the authors argue that these towns suffered similar damage so that their failure to receive an exemption was arbitrary. Cipollone and Rosolia find statistically significant higher graduation rates for females in the high schools subject to the exemption when compared to females in the comparable high schools that were not subject to exemption. One limitation of this type of calculation is that it is difficult to interpret in terms of social mechanisms, an issue recognized by the authors. Regardless of this, the finding itself may be problematic for reasons that are delineated in Heckman (1998). One reason involves what Heckman, Urzua, and Vytlacil (2006) have dubbed "essential heterogeneity": if the effect of increased male graduation on a girl's payoff differential between graduating and dropping out is heterogeneous, then it is unclear how to interpret the Cipollone and Rosolia conclusion that a 1% increase in male graduation leads to a 0.7% increase in female graduation in terms of policy counterfactuals, i.e., increases in male graduation rates need not have this effect on female graduation rates in schools that differ from those in the sample or contexts different from the particular quasi-experiment under study. Further, one can even question whether this statistical finding is evidence of a social interaction per se. Compulsory military service was previously subject to exemptions for high school graduation. Thus, the general exemption changed the composition of males in a school in particular ways. In essence the general exception affected the attendance of males whose unobservable characteristics made their graduation behavior especially sensitive to the policy change relative to the previous regime. Suppose that there is assortative matching on these unobservable characteristics in the formation of romantic relationships. One can reasonably imagine that an associated increase in graduation for females occurs because of the preservation of romantic relationships that would have been severed by school (and community) withdrawal for military service. The message of this possibility is that the translation of what amount to partial correlations on the behavior of one group with the behaviors of another

group into causal claims about social interactions that can answer policy relevant questions requires careful consideration of counterfactuals and the nature of unobservable individual-specific heterogeneity, which is a theme of Heckman's (1998) analysis.

b. Panel data

A second standard strategy for dealing with unobserved group interactions involves the use of panel data to difference the interactions out. Supposing that the variables in (30) are indexed by t, this amounts to working with

$$\begin{aligned}\omega_{igt} - \omega_{igt-1} = c(x_{it} - x_{it-1}) + d(\gamma_{gt} - \gamma_{gt-1}) \\ + J(m_{gt} - m_{gt-1}) + \varepsilon_{it} - \varepsilon_{it-1}.\end{aligned} \tag{32}$$

Recall that our identification theorem 1 depended on the relationship between $\bar{x}_g$, γ_g and m_g. For (32), theorem 1 immediately can be applied if one considers the requirements of the theorem as they apply to $\bar{x}_{gt} - \bar{x}_{gt-1}$, $\gamma_{gt} - \gamma_{gt-1}$ and $m_{gt} - m_{gt-1}$. So long as there is temporal variation in $\bar{x}_{gt}$ and γ_{gt} i.e., the first differences in (32) are not zero, then the conditions for identification will be the same as in the original linear model without α_g. Note that variation in $\bar{x}_{gt}$ and/or γ_{gt} will induce variation in m_{gt} over time. An early example of this strategy is Hoxby (2000) who focuses on variation in the percentage of a student's own ethnic group in a classroom.

For those elements of x_{it} and γ_{gt} that do not vary over time, differencing means that their associated coefficients will not be identified. Defining the time invariant elements of γ_{gt} as γ_g^1, the lack of identification of their associated parameters d^1 occurs for the obvious reason that one cannot differentiate the effect of $d^1\gamma_g^1$ from α_g. On the other hand, all elements of x_{it} may be identified if additional assumptions are placed on ε_{it}. As formally discussed in Graham and Hahn (2005), suppose that

$$\mathrm{E}(\varepsilon_{it}|F_{x|gt}, \gamma_{gt}, \alpha_{gt}, i \in g \text{ at time } t)) = 0.$$

In this case, intragroup variation in x_{it} at a single point in time can identify all of the elements of c. The reason for this is that for group g at a fixed t, $k + d\gamma_{gt} + Jm_{gt} + \alpha_{gt}$ acts as a constant term for the members of the group. Brock and Durlauf (2001b, 2006, 2007) use this same argument for cross-section identification of individual interactions coefficients in discrete choice models. As noted by Graham and Hahn, this type of argument is originally due to Hausman and Taylor (1981).

c. Variance approaches and group-level unobservables

Graham (2008) provides a strategy for identifying the parameter for endogenous interactions in the presence of unobserved group interactions in parallel to the arguments that identified J in (26). Graham works with the natural generalization of (26):

$$\begin{aligned}\omega_{it} &= J\bar{\omega}_g + \alpha_g + \varepsilon_i \\ &= Jm_g + \alpha_g + \varepsilon_i + J\bar{\varepsilon}_g.\end{aligned}$$

Critically, Graham assumes that α_g is a random effect, specifically requiring that the conditions

$$\operatorname{cov}(\alpha_g \varepsilon_i | i \in g) = 0 \tag{33}$$

and

$$\operatorname{var}(\alpha_g | i \in g) = \sigma_\alpha^2 \tag{34}$$

hold in addition to equations (27) and (28). Equation (33) states that individual and group unobservables are uncorrelated and is justified in Graham's context by the random assignment of teachers across classrooms. Equation (34) rules out any dependence of the variance of unobserved group effect on group size. In a classroom context, this means that the variance of teacher quality does not depend on the number of students. In a direct generalization of equation (29), Graham (2008) shows that

$$\operatorname{var}(\omega_g) = \left(I_{n_g} - \frac{J}{n_g}\iota_{n_g}\right)^{-1} (\sigma_\alpha^2 \iota_{n_g} + \sigma_\varepsilon^2 I_{n_g}) \left(I_{n_g} - \frac{J}{n_g}\iota_{n_g}\right)^{-1}, \tag{35}$$

which means that one can again use differences in the variance of outcomes across groups of different sizes to identify J.

Taken together, assumptions (27) and (34) are, in our judgment, problematic for an experiment designed to assess how classroom size affects learning; while the objective of the experiment was presumably to understand how first moments are affected, there is no reason to think that second moments would be unaffected if the first moments are classroom-size dependent. In other words, it is one thing to argue that prior to assignment, students are exchangeable with one another and teachers are exchangeable with one another. After assignment, one of the student's characteristics is the size of the classroom to which he has been assigned. Effects of classroom size on the variance of shocks is especially plausible if the mechanisms that produce teacher quality and idiosyncratic student outcomes are nonlinear, which is a major issue in the econometrics of discrimination.[23] That said, Graham's recognition that the assumption that if unobserved group effects are random rather than fixed leads to new routes to identification is a significant methodological advance. And of course, our point is not that Graham's empirical claims are incorrect, but rather that a reader needs to assess the plausibility of the assumptions that permit identification for his context.

v. Self-selection

It is natural for many social contexts to expect individuals to self-select into groups. This is most obvious for the case of residential neighborhoods; models such as Bénabou

[23] See Heckman (1998).

(1993, 1996), Durlauf (1996a,b) and Hoff and Sen (2005), for example, all link social interactions to neighborhood choice. In terms of estimation, self-selection generally means that equation (8) is violated.

Self-selection has typically been addressed using instrumental variables methods. Evans, Oates, and Schwab (1992) is an early example. The use of instrumental variables as a solution to self-selection suffers, in our view, from the problem of theory openendedness as was discussed in the context of unobserved group effects. However, unlike the case of unobservable group interactions, self-selection involves a specific behavior on the part of the agents under study which can provide additional insight into instrument validity. For example, Evans, Oates, and Schwab (1992) focus on estimating the effect of the percentage of students in a school who are disadvantaged on high school dropout and teen fertility rates. The measure of school level socioeconomic disadvantage is instrumented with metropolitan area levels of unemployment, college completion, poverty rates and median income. The instruments are justified on the grounds that while families may choose schools within a metropolitan area, they are unlikely to choose metropolitan areas because of schools. This may be correct as far as it goes, but the relevant question for instrument validity is whether the instruments are uncorrelated with ε_i. One obvious reason why this is true is that drop out and pregnancy decisions will be related to labor market opportunities, which by the logic of Evans, Oates and Schwab's choice of the instruments would be defined at the metropolitan and not the school level. Durlauf (2004), on the other hand, suggests reasons why the instruments may not be valid.

A distinct source of instruments to account for self-selection has been quasi-experiments. A recent example is Lalive and Cattaneo (2009) who study the effect on schooling decisions in Mexico by the PROGRESA program. This program rewards school attendance by rural children with small cash subsidies to their families. Lalive and Cattaneo measure social interactions via the effects by schoolmates whose families are eligible for PROGRESA on their ineligible classmates. The basic idea is that if PROGRESA directly raised attendance of the first set of children, peer effects must explain why attendance among the ineligible increased. This type of exercise is informative, but again suffers from interpretation problems of the type we have raised in the context of Cipollone and Rosolia (2007). We simply note here that the Heckman (1997) and associated critiques have equal force when instrumental variables are employed to account for self-selection as when they are used to account for group level unobservables.

In our view, the preferred approach to dealing with self-selection is to treat group choice and behavior within a group as a set of joint outcomes, and conduct empirical analysis from the perspective of both behaviors. Unlike the instrumental variables approach, this has interesting implications for identification, at least for the linear model. Brock and Durlauf (2001b) first recognized this possibility and studied the case of self-selection between two groups; Brock and Durlauf (2002, 2006) and Ioannides

and Zabel (2008) extended this analysis to an arbitrary finite number of groups. At an intuitive level, this is not surprising. Self-selection represents a behavior on the part of an agent and so should contain information about his preferences, which will depend on the social interactions that occur in groups over which he is choosing.[24] Unlike the instrumental variable approach, modeling self-selection exploits this information rather than treats it as a nuisance.

Following Heckman's original (1979) reasoning, one can think of individuals choosing between groups $g = 1,\ldots,G$ based on an overall individual-specific quality measure for each group:

$$I_{ig}^{*} = \gamma_1 x_i + \gamma_2 \gamma_g + \gamma_3 z_{ig} + \nu_{ig},$$

where z_{ig} denotes those observable characteristics that influence i's evaluation of group g but are not direct determinants of ω_i and ν_{ig} denotes an unobservable individual-specific group quality term. Individual i chooses the group with the highest I_{ig}^{*}. We assume that prior to group formation, for all i and g, $\mathrm{E}(\varepsilon_i|x_i, \gamma_g, z_{ig}) = 0$ and $\mathrm{E}(\nu_{ig}|x_i, \gamma_g, z_{ig}) = 0$.

From this vantage point, the violation of equation (8) amounts to

$$\mathrm{E}(\varepsilon_i|x_i, \bar{x}_1, \gamma_1, z_{i1}, \ldots, \bar{x}_G, \gamma_G, z_{iG}, i \in g) \neq 0. \tag{36}$$

Notice that equation (36) includes the characteristics of all groups. This conditioning reflects the fact that the choice of group depends on characteristics of the groups that were not chosen in addition to the characteristics of the group that was chosen. Equation (36) suggests that the linear in means model, under self-selection, should be written as

$$\omega_{ig} = cx_i + d\gamma_g + Jm_g + \mathrm{E}(\varepsilon_i|x_i, \bar{x}_1, \gamma_1, z_{i1}, \ldots, \bar{x}_G, \gamma_G, z_{iG}, i \in g) + \xi_i, \tag{37}$$

where by construction $\mathrm{E}(\xi_i|x_i, \bar{x}_1, \gamma_1, z_{i1},\ldots, \bar{x}_G, \gamma_G, z_{iG}, i \in g) = 0$. Notice that the conditioning in (36) includes the characteristics of all groups in the choice set. This is natural since the characteristics of those groups not chosen are informative about the errors.

Equations (36) and (37) illustrate Heckman's (1979) insight that in the presence of self-selection on unobservables, the regression residual ε_i no longer has a conditional mean of zero, yet (37) can be consistently estimated using ordinary least squares if one adds a term to the original linear in means model (6) that is proportional to the conditional expectation on the left hand side of (36), i.e., prior to estimation. Denote this estimate as

$$\overbrace{\kappa \mathrm{E}(\varepsilon_i|x_i, \bar{x}_1, \gamma_1, z_{i1}, \ldots, \bar{x}_G, \gamma_G, z_{iG}, i \in g)}. \tag{38}$$

[24] The idea that full consideration of the informational contents of choices can be exploited to overcome issues of unobservable heterogeneity and generate identification appears in other contexts. In Heckman and Honoré (1990), identification of the skill distribution in a population depends on cross market variation in skill premia, which bears a strong analogy to our use of cross-group variation in selection corrections.

Heckman's fundamental insight was that one can construct such a term by explicitly modeling the choice of group. From this perspective, controlling for self-selection amounts to estimating

$$\omega_{ig} = cx_i + dy_g + Jm_g + \overbrace{\rho\kappa \mathrm{E}(\varepsilon_i | x_i, \bar{x}_1, y_1, z_{i1}, \ldots, \bar{x}_G, y_G, z_{iG}, i \in g)} + \xi_i. \tag{39}$$

Thus, accounting for self-selection necessitates considering identification for this regression, as opposed to (6).

The property of interest for the identification of social interactions is that the addition of the term (39) can help facilitate identification. To see this, consider two possible reasons why agents choose particular groups. First, agents may choose groups on the basis of the expected average behaviors that occur. For example a family chooses a neighborhood based on its expectation of the average test score among students in the school their child will attend. In the extreme case where this is the only neighborhood factor that matters to families, the conditional expectation associated with the selection correction will be a function of the agent's characteristics and the expected outcomes in each of the neighborhoods, i.e.,

$$\mathrm{E}(\varepsilon_i | x_i, \bar{x}_1, y_1, z_{i1}, \ldots, \bar{x}_G, y_G, z_{iG}, i \in g) = \varpi(x_i, m_1, \ldots, m_G). \tag{40}$$

By the same logic that rendered the partial linear model (15) identified, (39) is also identified as m_g cannot, outside of nongeneric cases, be linearly dependent on a constant term and y_g.

Second, parents may choose neighborhoods based on the mean incomes of families or some measure of the distribution of occupations among neighborhood adults. This can be justified on role model grounds. If neighborhoods are evaluated according to their contextual variables, then (38) functions as an additional individual-specific regressor whose group level average does not appear in (37). Hence, following the argument about identification in linear in means models that was developed earlier, the presence of a regressor with a nonzero coefficient can allow for identification to occur. This route to identification has been successfully used in Ioannides and Zabel (2008) to identify social interactions in housing. See Ioannides (this volume) for more detail on that study.

One limitation of the work that has been done on identification using selection corrections is that it has employed parametric corrections. Brock and Durlauf (2002, 2006), for example, work with Lee's (1983) correction which assumes joint normality of ε_i and v_{ig}, while Ioannides and Zabel (2008) work with joint extreme value distributions, following Dubin and McFadden (1984). From the perspective of econometric theory, selection corrections are now conceptualized as examples of control functions, i.e., functions which augment models in order to account for endogeneity as well as self-selection. Navarro's (2008) survey highlights advances made in the nonparametric

construction of control functions. This suggests that the argument we have made about the ability of self-selection to facilitate identification of social interactions can move beyond parametric assumptions, although this has not yet been done. We believe this is feasible given the parametric structure, i.e., the partial linearity of the outcome equation under self-selection.

vi. Social interactions via unobserved variables

Our discussion of the linear in means model has assumed that the variables through which social interactions operate either are directly observable or represent rational expectations forecasts of observable (to the analyst) variables. Recent work by Arcidiacono, Foster, Goodpaster, and Kinsler (2009) considers this possibility in a panel context, which is applied to classrooms at the University of Maryland, where grades are the outcome measure. Translating their model into our notation, they analyze

$$\omega_{igt} = cx_i + d\bar{x}_{gt} + eu_i + f\bar{u}_{gt} + \varepsilon_{it}, \tag{41}$$

where $\gamma_{gt} = \bar{x}_{gt}$ is assumed. Endogenous social interactions are ruled out *a priori*. The key innovation in Arcidiacono *et al.* is that the individual variable u_i and associated group variables $\bar{u}_{gt}$ are both unobservable. Notice that x_i and u_i are time invariant whereas $\bar{u}_{gt}$ and $\bar{x}_{gt}$ are time dependent. The time dependence of the latter terms occurs because group memberships can change over time.

The identification of social interactions for this problem thus hinges upon overcoming unobservability of the contextual interactions $\bar{u}_{gt}$. In order to achieve identification, Arcidiacono *et al.* restrict the coefficients in (41) by assuming the existence of a scalar γ such that

$$e = \gamma c; \quad f = \gamma d.$$

This assumption follows Altonji, Huang, and Taber (2005). In the spirit of simultaneous equations theory this is analogous to a coefficient restriction that facilitates identification. This assumption fixes the ratios of the coefficients of observed individual characteristics to equal those of the corresponding peer characteristics. Arcidiacono *et al.* describe this (p. 6) as rendering the two dimensions of peer effects versus the two dimensions of individual effects "equally important". This is an ill-defined claim. The Arcidiacono *et al.* strategy is better thought of as a restriction on coefficients that, in the classical simultaneous equations sense, may help with identification and its justification should be assessed from the perspective of whether the restriction can be justified by economic theory or by some other argument.

Arcidiacano *et al.* proceed by focusing on a general notion of fixed effects for each individual, defining these fixed effects the determinants of outcomes (outside the errors ε_{it}) as

$$\kappa_i = cx_i + eu_i.$$

Letting $\bar{\kappa}_{gt} = \frac{1}{n_{gt}} \Sigma_{i \in g} \kappa_{igt}$, equation (41) can be re-expressed as

$$\omega_{ig} = \kappa_i + \gamma \bar{\kappa}_{gt} + \varepsilon_{it}.$$

From this perspective, Arcidiacono *et al.* frame the identification problem for social interactions as the problem of consistently estimating γ in the presence of a large number of fixed interactions. Their theorem 1 locates a set of sufficient conditions so that a consistent and asymptotically normal estimator of γ may be found. Identification is implicit in this proof. We do not repeat their assumptions here but note that the essential substantive economic requirements are 1) the composition of an individual's peer groups change over time and 2) $\forall i, j, t$, $\mathrm{E}(\varepsilon_{it}\kappa_j) = 0$. The first condition is needed since identification requires individuals to be exposed to different peers to allow for distinguishing the influence of the fixed effects of others on a given individual. The second condition delimits the nature of self-selection into classrooms. Although Arcidiacono *et al.* argue that they allow for self-selection based on the ability of peers, this second condition appears to limit how selection can occur. This criticism does not detract from the value of their contribution, but points to an instance of the general proposition that explicit modeling of selection is essential in understanding identification conditions.

Cooley (2009) explores an even more extreme form of unobservability in which both the determinants and outcomes of behaviors are not observed. Motivated by the issue of classroom peer interactions, she argues that, from the vantage point of theory, the causal sources of peer interactions between students in a classroom involve the effort levels of other students, which is an endogenous effect and ability, which is a contextual effect. Neither of these is directly observable. Empirical work on classroom peer interactions typically uses observable classroom achievement as the outcome measure and various ad hoc (albeit theoretically plausible) choices of observable student characteristics for contextual efforts. For Cooley, what is critical is that effort is the choice variable, not classroom performance per se, which will reflect unobservable effort and ability as well as residual unobserved heterogeneity. Cooley calculates the optimal choice of effort by students and uses this equilibrium condition to produce a reduced form model, which links the observable outcome classroom performance to the observable individual and contextual interactions. The coefficients in this latter equation prove to be uninterpretable. For example, an observable contextual effect whose structural coefficient is positive can have a reduced form coefficient that is negative. The formulas linking the structural and reduced form parameters are extremely complicated, so are omitted here.

A third approach to unobservability is developed in Solon, Page, and Duncan (2000) and further studied in Page and Solon (2003a,b). This analysis assumes that one cannot observe any of the determinants of individual outcomes; only the outcome data are available. Unlike work such as Graham (2008), no assumption is made that social interactions are endogenous rather than contextual. Rather, it is assumed that individuals are influenced

by family-level, group-level, and idiosyncratic influences. Individuals are distinguished by family and group (in this case residential neighborhood). A variance decomposition for individual outcomes is constructed to bound the contribution of group effects to variance of individual outcomes. The analysis may be understood in terms of a variance components model (Searle, Casella, and McCulloch, 2006, p. 14):

$$\omega_{ig} = \mu_f + \nu_g + o_{fg} + \varepsilon_i. \tag{42}$$

In (42), μ_f denotes a family effect, ν_g denotes a group effect, o_{fg} denotes an interactive effect between family and group, and ε_i denotes an idiosyncratic effect. A decomposition of this type always exists in which the components are orthogonal. In terms of mapping this expression back to a measure of the role of group influences, one difficulty lies with o_{fg}. Does the covariation between family and group represent a group effect or self-selection? Solon, Page and Duncan address this issue by comparing intra-family (sibling) and group (in their case residential neighborhood) variances to bound the variance of ν_g, finding the variance contribution is small. Oreopoulos (2003) finds similar results, focusing on a data set that involves adults who, as children, were randomly assigned to different public housing projects, thereby presumably eliminating o_{fg}. He finds a small role for ν_g and so concludes that neighborhood effects do not play a major role in explaining the variance of various adult outcomes.

One limitation of this approach is that it reduces the vector of social interactions to a scalar so that one cannot tell which social factors matter. In fact, it is possible for different social factors to cancel each other out. And to the extent that unobserved group effects ν_g do not represent social interactions, as was the case for our example of teacher quality and classroom outcomes, it is not clear that a large variance contribution from ν_g has a social interactions interpretation. These caveats do not render such exercises uninteresting; rather they illustrate how economically substantive assumptions matter in producing economically substantive interpretations.

vii. Social multipliers and information from aggregated data

We close this section on the linear in means model by turning to how the relationship between individual and aggregated data may provide evidence of social interactions when individual actions are generated by the linear in means model. The relationship between individual and aggregated data is studied in Glaeser and Scheinkman (2002) and applied by Glaeser, Sacerdote, and Scheinkman (2003). The essential idea behind their analysis is that endogenous social interactions can generate social multipliers in the sense that a change in private incentives for every agent in a population will have an equilibrium effect that is greater than the direct effect of the incentive change on each individual because the changes in the behavior of others create additional effects on that individual. This difference is evident in the reduced form (11) since x_i and $\bar{x}_g$ have different coefficients.

Focusing on a scalar case (extension to vectors is straightforward but algebraically tedious), Glaeser and Scheinkman (2002) propose comparing the coefficient b in the regression

$$\omega_{ig} = a + bx_i + \varepsilon_i$$

with the coefficient b' in its group counterpart

$$\bar{\omega}_g = a' + b'\bar{x}_g + \bar{\varepsilon}_g, \tag{43}$$

and define the social multiplier as the coefficient ratio:

$$S = \frac{b'}{b}.$$

In the context of the linear in means model, it is straightforward to compute this ratio. Our calculations differ from Glaeser and Scheinkman as we focus on the difference in the regressions as a misspecification problem. To make this calculation, notice that for the Glaeser and Scheinkman case, the reduced form (11) becomes:

$$\omega_{ig} = \frac{k}{1-J} + cx_i + \frac{cJ}{1-J}\bar{x}_g + \varepsilon_i. \tag{44}$$

Comparing (44) and (43), it is evident that the population value of b is readily calculated using the standard omitted variables formula that $b = c + \frac{cJ}{1-J}\beta$ where β is implicitly defined by $\text{proj}(\bar{x}_g \,|1, x_i) = \kappa + \beta x_i$, i.e., $\beta = \text{cov}(x_i, \bar{x}_g)/\text{var}(x_i)$. In contrast, it is evident from taking expected values on both sides of (44) that $b' = \frac{c}{1-J}$ which means the social multiplier is

$$S = \frac{1}{1 - J + J\beta}. \tag{45}$$

Notice that if there is perfect segregation across groups, so that incomes within a group are identical, then $\text{cov}(x_i, \bar{x}_g) = \text{var}(x_i)$, which implies that $S = 1$ whereas under random assignment, $\text{cov}(x_i, \bar{x}_g) = 0$, which implies $S = \frac{1}{1-J}$. The latter value of S takes a form that echoes the classic Keynesian income/expenditure multiplier, with the marginal propensity of consumption replacing J. Moreover, in the Bayesian game described in appendix 1, which lays out the underlying decision-theoretic framework of the linear in means model, $S = (1 + \phi)/(1 + \phi\beta)$ where ϕ measures the marginal rate of substitution between conformity and the private value of the choice variable. Note the surprising fact that, as the pressure to conform increases, the social multiplier may either decrease or increase, depending upon whether β exceeds or is exceeded by 1. Both are possible since we have not specified how individuals are sampled across groups. For example, if there is systematic sampling of x_i values below $\bar{x}_g$ in each group, then $\beta > 1$ may occur. The dependence of the social multiplier upon β makes it difficult to interpret it.

As articulated in this example, the social multiplier provides a different perspective on the effects of endogenous social interactions on changes in private incentives. In terms of identification, it may also be of value. Clearly, the social multiplier calculations have little to add if complete individual level data are available across the various groups of interest. On the other hand, suppose that aggregate data are incomplete; i.e., one knows about outcomes in a subset of classrooms in a school. One can imagine identification of J via analogs to (45) that compare different levels of aggregation and thus exploit variation in β to uncover J. Alternatively, one can imagine partial identification approaches that exploit the fact that different β's reflect different levels of aggregation with respect to the same underlying population.

4. SOCIAL NETWORKS AND SPATIAL MODELS OF SOCIAL INTERACTIONS

In defining social interactions thus far we have presumed that interactions are generated by group-specific averages. Social network models provide further focus on the microstructure of interactions among agents and allow for heterogeneity of interactions across pairs of agents. Jackson (2008) provides a thorough overview of the new social networks literature. In this section we address the identification of social interactions in social networks. In addition, we discuss the use of spatial econometrics models to study social interactions. The social networks and spatial analysis approaches are mathematically very similar, and yet, they have been until recently developed independently from one another. This similarity is not surprising as spatial econometrics approaches deal with physical space, whereas social networks address a more abstract social space, yet still a space with well posed notions of distance and the like.

i. Graphical models of social networks

Before extending the model of social interactions to social networks it is useful to establish some basic terminology. For this setting, social interactions among individuals are defined by means of a social structure (or topology, the two terms are used interchangeably in the literature and here) that takes the form of a network, whose mathematical description is a graph with the vertices representing individuals and edges representing links between them. Network vertices and population members are thus identical concepts. What is of interest is the network structure that links agents. Network structure among individuals is modeled either as an undirected or a directed graph. Here we shall focus on the latter case, since it allows us to express a richer set of social relations.

Directed graphs consist of vertices (also known as nodes) and directed edges. A directed edge is an ordered pair (i, j) of vertices. A directed graph is a pair (V, E) where V is the set of nodes, with cardinality n_V, and E is the set of edges. A subgraph (V', E') of (V, E) is a graph where V' is a subset of V and E' is a subset of E. A subgraph (V', E')

is induced by (V, E) if and only if E' contains all edges of E which begin and end in V'. A social network is a graph (V, E) where V is the set of individuals and the directed edges in E signify social influence: (i, j) is in E if and only if j influences i.

A social network can be represented by its adjacency matrix A, also known as its sociomatrix in the mathematical sociology literature. An adjacency matrix is an $n_V \times n_V$ matrix, with one row and one column for each individual in V. For each pair of individuals i and j, $a_{ij} = 1$ if there is an edge from i to j, and 0 otherwise. Since the network is supposed to represent social connections, it is natural to assume that no i is connected to himself. That is, for all i, $a_{ii} = 0$. A path from i to j is a sequence of individuals $i_0,\ldots, i_K$ such that $i_0 = i$, $i_K = j$, and for all $k = 1,\ldots,K - 1$, there is an edge from i_{k-1} to i_k. Such a path is said to have length K. If there is a path from i to j of length exceeding 1, then j indirectly influences i. The adjacency matrix A displays all paths of length 1. The K-fold product A^K counts all paths of length K; if the ij'th element of A^K is n, then there are n paths of length K from i to j. A subgraph (W, F) of (V, E) is strongly connected if and only if for any i and j in W there is a path from i to j consisting solely of edges in F. A subgraph which is strongly connected and is a subgraph of no larger such graph is a strongly connected component of (V, E).[25] A graph (V, E) is strongly connected if and only if some power of its adjacency matrix is strictly positive. The literature also contains the less restrictive requirement of weak connectivity. Intuitively, it is the notion of connectivity that emerges when one can walk links in any direction. By suitably ordering the vertices, the adjacency matrix of a graph (V, E) can be written as a block-diagonal matrix where the rows (columns) of each block correspond to a weakly connected component.[26] A graph (V, E) is complete if for each pair i and j in V there is an edge from i to j. A graph is directed if the existence of an edge from i to j implies that there is no edge from j back to i.

While social influence can be a one-way relationship, we usually think of some relationships, for example friendship, as being bidirectional and the social network is represented by an undirected graph, and the adjacency matrix is symmetric. Edges are now undirected, and so there is a path from i to j if and only if there is a path from j to i. A subgraph (V', E') of the undirected graph (V, E) is connected if and only if between any two nodes i and j in V' there is a path in (V', E') between them. A component of the graph (V, E) is, as before, a subgraph which is connected and maximal with respect to inclusion. The distance between any two nodes is the length of the shortest path between them.

Some particular network topologies are important in the social networks literature. A *star network* is an undirected graph in which one individual, the center, is connected

[25] Note that a strongly connected component can have links into it from outside the component, and links to the outside. But no other node is on a path *from* the component *and* a path *to* the component.

[26] Weak connectivity is connectivity assuming that all edges can be traversed in any direction.

to all other individuals while all other individuals are connected only to the center. A *group*, also known as a *complete network*, is one that contains an edge between each two of its vertices. Its adjacency matrix is all 1s. In a *bipartite graph*, the vertex set V is the union of two disjoint sets T and U, and all edges are between members of T and members of U, i.e., edges represent matches between vertices in the two sets. We will call a bipartite graph *directed* if for all $(i, j) \in E$, $i \in T$ and $j \in U$.

Sociologists allege that social relations like friendship exhibit the property of homophily — loosely but accurately described by the phrase "the friend of my friend is my friend, too." This property is described by the prevalence of transitive triads. Triads are connected subgraphs consisting of three nodes. Transitivity is the property that the existence of an edge from node i to j and an edge from j to k implies the existence of an edge from i to k. A graph is transitive if it contains no intransitive triads. The linear in means model is specified by assuming A is symmetric, that edges are bidirectional, and that the graph is transitive. If this is true, then the graph is the union of completely connected components. The nodes of the component containing i constitute i's group.

While the linear in means model is a good starting point for the study of social interactions, social networks allow for a much richer specification of social relations. The model can be enriched still further by allowing the elements of adjacency matrices to be arbitrary real numbers. In such models, the magnitude of the number a_{ij} measures the degree of influence j has on i, and the sign expresses whether that influence is positive or negative. Throughout this section we will assume that all elements a_{ij} are non-negative except as noted, and that contextual variables are weighted averages of the individual characteristics. This generalizes the contextual effects in the linear in means model case in which $\gamma_g = \bar{x}_g$.

Note that in this section, we defined choices as ω_i. No additional subscript is employed to denote a person's social environment. The reason for this is that the networks literature focuses on members of a common population and introduces social structure for the population as a whole via the matrix A.

ii. Identification in social networks: basic results

Cohen-Cole (2006) appears to be the first analysis of the linear in means model under richer interactions structures. That is, he posits that an individual reacts to multiple reference groups, such as a teenage boy might care differently about what other teenage boys do than about what teenage girls do. He shows that the model with agents' beliefs about actions in multiple other groups as well as observables for each group can be fully identified provided that there are more observed linearly independent group level effects than there are groups in the sample and that there is some pair of groups for which agents in one group care about the actions in the other. This type of reasoning is extended in an analysis by De Giorgi, Pellizari, and Redaelli (2010) of peer effects in the choice of college education. From the econometric perspective, Bramoullé, Djebbari, and Fortin (2009), Lee, Liu, and Lin (2010) and Lin (forthcoming) constitute

the most systematic explorations of social interactions in social and spatial contexts respectively, but there are several other contributions which we will discuss below.

The current identification literature for networks assumes A is known *a priori* to the researcher. This is a critical assumption in the existing literature that restricts empirical work to contexts in which survey data, for example, can be used to measure network structure. As far as we know, no econometric methods have been developed to consider social interaction inferences in the presence of unknown network structure. We take this up in section 4.vii below. We note that some work has focused on data sets in which a random sample is drawn from a set of networks. We ignore the distinction between possessing the data on the whole population versus a random sample as it is not germane to the identification issues on which we focus.

The network model employed by Bramoullé *et al.* assumes that each individual i is influenced by the average behavior of a set of peers $P(i)$ and that, like Moffitt (2001) but unlike Manski (1993), individual i is not his own peer. The peer relationship is not assumed to be symmetric, so the social network is represented by a directed graph. The social interactions are described by a weighted adjacency matrix:

$$a_{ij} = \begin{cases} \dfrac{1}{|P(i)|} & \text{if } j \in P(i), \\ 0 & \text{otherwise.} \end{cases} \tag{46}$$

Individual outcomes are then described by the behavioral equation system

$$\omega_i = k + cx_i + d\sum_{j \neq i} a_{ij}x_j + J\sum_j a_{ij}\omega_j + \varepsilon_i \tag{47}$$

with the error restriction

$$\mathrm{E}(\varepsilon_i | (x_i)_{i \in V}, A) = 0. \tag{48}$$

The reduced form for this system may be described in vector notation as

$$\omega = k(I - JA)^{-1}\iota + (I - JA)^{-1}(cI + dA)x + (I - JA)^{-1}\varepsilon, \tag{49}$$

where I refers to the $n_V \times n_V$ identity matrix and ι is a $n_V \times 1$ vector of 1's. (Recall that n_V is the number of individuals in the network.) Bramoullé *et al.* focus on identification by studying this reduced form. Recognizing, as did Moffitt (2001), that systems of this type are examples of linear simultaneous equations models in which one can think of the outcomes for the members of the overall network as the endogenous variables and the individual effects as the exogenous variables.[27] The important insight of Bramoullé *et al.* is that the ideas concerning the averages of behaviors and characteristics of groups carry over into more general social-network settings. Bramoullé *et al.* provide

[27] If γ_g were included in the system (47), this vector would also represent a set of exogenous variables.

a fundamental algebraic result with respect to identification of models like equations (47) and (48), which does not rely on the constraint (46). The theorem assumes that J can take values in an arbitrary parameter set $\mathcal{J}$ in $\mathbf{R}$.

Theorem 2. Identification of social interactions in linear network models. *For the social interactions model described by equations* (47), *and* (48), *assume that* $Jc + d \neq 0$ *and that for all values of* $J \in \mathcal{J}$, $(I - JA)^{-1}$ *exists.*

i. *If the matrices* I, A, *and* A^2 *are linearly independent, then the parameters* k, c, d *and* J *are identified.*

ii. *If the matrices* I, A, *and* A^2 *are linearly dependent, if for all* i *and* j, $\sum_k a_{ik} = \sum_k a_{jk}$, *and if either* A *has no row in which all entries are 0 or* $k = 0$, *then parameters* k, c, d *and* J *are not identified.*

The condition that $Jc + d \neq 0$ requires, in the network setting, that endogenous and contextual effects do not cancel out in the reduced form. Theorem 2.i is a purely algebraic result. This is to say, it does not rely on the specific structure of A, which arises from its network context. It applies to any linear system of the form (47) for which $|J| \cdot \|A\| < 1$ for all possible parameter values J. An interesting feature of this result is that it does not rely on exclusion restrictions. This should not be surprising. Although the number of equations in the system is n_V, the size of the population, there are only four parameters to estimate. There are thus many cross-equation and within-equation linear equality constraints: The independence condition describes when these constraints satisfy the appropriate rank condition. Theorem 2.ii identifies an important case for which linear independence of I, A and A^2 is necessary as well as sufficient for identification. The requirements on A mean that each individual averages in some way over those who influence him, and that no one is isolated. The result can be thought of as a converse to theorem 2.i.

The analysis of group interaction is the leading case in the econometric literature on networks. It is also appealing from the perspective of existing data sets such as the National Longitudinal Study of Adolescent Health (Add Health).[28] Suppose that the peer relation is symmetric, $j \in P(i)$ if and only if $i \in P(j)$. Suppose too that the peer relation is transitive: If $j \in P(i)$ and $k \in P(j)$, then $k \in P(i)$. As discussed in section 4.i, the graph is the union of a finite number G of completely connected components, that is, groups. Suppose that component g has n_g members. We will consider two ways to average over the group: *Exclusive averaging* excludes i from $P(i)$. In this case, for $i \in g$,

$$a_{ij} = \begin{cases} \dfrac{1}{n_g - 1} & \text{if } j \neq i \text{ and } j \in g, \\ 0 & \text{otherwise.} \end{cases}$$

[28] The Add Health data set is the outcome of a longitudinal data collection exercise designed to facilitate study of health-related behaviors of adolescents in grades 7 through 12. The data set includes information, for example, on the structure of adolescent friendships via responses to questions on the identities of best friends. As such, directions of friendships are revealed but not intensity. See http://www.cpc.unc.edu/projects/addhealth for details.

Inclusive averaging includes i in $P(i)$. In this case, for $i \in g$,

$$a_{ij} = \begin{cases} \dfrac{1}{n_g} & \text{for all } j \in g, \\ 0 & \text{otherwise.} \end{cases}$$

With inclusive averaging, equations (47) and (48) are equivalent to our linear in means model, except that realized rather than expected outcomes affect individual outcomes. This difference is inessential since the instrumental variable projections used to replace the endogenous choices of others coincide with equilibrium formulations of beliefs.[29] Means and realizations, however, represent two distinct theoretical models. The first is a network version of the incomplete-information game developed in appendix 1. The second is a complete-information version of the same game. With exclusive averaging, the subject of Bramoullé *et al.*, an additional distinction is that the calculation of group-level contextual effects does not include i's own individual characteristics. This distinction is inessential in that the identification results for the two models are nearly the same. The following result is a corollary of the forthcoming theorem 3. It highlights the one case where the two averaging techniques differ:

Corollary 1. Identification of social interactions in group structures with different-sized groups. *Suppose that individuals act in groups, and that the a_{ij} are given by exclusive averaging. Then the parameters k, c, d and J are identified if there are at least two groups of different sizes. With inclusive averaging, the parameters are not identified.*

The positive result of corollary 1 is similar to Graham's (2008) variance contrast identification strategy, but its source is different. Here identification follows the reduced form regression parameters rather than the second moments of the average group outcomes.[30] Note that in Graham's case, $Jc + d = 0$ since $c = d = 0$, so his findings allow for identification when individual and contextual effects are absent.

At first glance, this corollary might appear to contradict theorem 1 and indeed call into question Manski's nonidentification results on the linear in means model, since neither involved groups sizes while the corollary links groups sizes to identification.

[29] As we will see, the distinction is important for binary choice models.

[30] Corollary 1 is a special case of Lee's (2007) result, without fixed effects. Lee (2007) studied the effects of group size on identification while also allowing for unobserved group fixed effects. He establishes identification of both the endogenous and exogenous social interactions provided there are sufficient variations in group sizes, but under somewhat restricted conditions relating group sizes to the total number of observations, and also provides asymptotic estimation properties. Davezies, d'Haultfoeuille, and Foughre (2009) show that the identification results hold under different conditions than Lee's. For Davezies *et al.*, identification holds so long as 1) the sizes of groups do not depend on overall sample size, as in Lee (2007), and 2) there exist at least three different group sizes. This avoids Lee's requirements that group sizes are linked to population size. Intuitively, this is possible because variations in group sizes create variations in reduced form coefficients across groups. This variation is evident from an example like that in Bramoullé, Djebbari, and Fortin (2009, p. 49), where the reduced form coefficients depend on the size of the group to which an individual belongs.

In fact, there is no contradiction. Theorem 1 and Manski's earlier analysis did not treat social interactions in the linear in means model as a simultaneous equations system that explicitly relates individual choices to one another within a group. More generally, previous studies of identification of the linear in means model have taken the effects of group averages as the objects of interest, not the pairs of cross-individual effects. In contrast, the linear in means model as appears in the econometrics literature is a large sample approximation to the solution of a particular Bayes/Nash game, as shown in appendix 1.[31] If one relaxes the approximation, then the coefficients of the linear in means model as it applies to a given group depend on the group's size. When groups sizes differ, the coefficients of their associated linear in means representations differ. When coefficients from groups of different sizes are combined, this allows one to uncover the parameters k, c, d and J.

Powers of A describe the network topology. When we examine this in detail, we find that for very few networks are parameter estimates not identified in the reduced form. Bramoullé *et al.* provide an algebraic condition that is necessary and sufficient for identification in a class of linear social interaction models. We show that identification fails only when the network is the union of groups. It is ironic that the importance of groups is due to their prevalence as a common specification in econometric studies, for it is only with groups that the identification issue even arises. We state and prove (in appendix 2) a variant and unification of the Bramoullé *et al.* results.

Theorem 3. Nonidentification of social interactions in network models under exclusive and inclusive averaging. *Assume* $Jc + d \neq 0$, $k \neq 0$ *and that for all values of* J, $(I - JA)^{-1}$ *exists.*

i. *Under exclusive averaging, the parameters* k, c, d *and* J *are not identified in the reduced form* (49) *for the structural model* (47) *and* (48) *if and only if the social network* (V, E) *is the union of weakly connected components wherein each component is a group, and all groups are the same size.*

ii. *Under inclusive averaging, the parameters* k, c, d *and* J *are not identified in the reduced form* (49) *for the structural model* (47) *and* (48) *if and only if the social network* (V, E) *is the union of groups.*

This theorem depends upon the particular weighting schemes used to average across peers. It follows from the more general theorem 4, which is interesting in its own right because it applies to any weighting scheme which puts positive weight on all edges of E.[32]

Theorem 4. Nonidentification for weighted averaging implies network transitivity. *Let* (V, E) *be a network with weighted adjacency matrix* A *as described by* (46). *Assume that* $Jc + d \neq 0$ *and that for all values of* J, $(I - JA)^{-1}$ *exists. If the parameters*

[31] Compare equations (85) and (86) in appendix 1.

[32] Bramoullé *et al.* discuss results of this type, but provide no general theorem like this.

k, c, d and J are not identified, then (V, E) is transitive. If the network is undirected, then (V, E) is the union of groups.

Finally, we note that all of this is based on the fundamental independence criterion of theorem 2, which applies to any matrix A no matter what its source, so long as it satisfies an algebraic criterion. From this general point of view, it is clear that nonidentification of parameters in the reduced form is rare. We suppose without loss of generality that the parameter J takes values $[0, 1)$, and denote by S the set of all matrices A such that $(I - JA)$ is invertible. If the matrices are $n_V \times n_V$, S is a semi-algebraic set of full dimension in $\mathbf{R}^{n_v^2}$.[33]

Theorem 5. Generic identifiability of the linear social networks model. *The set of all matrices $A \in S$ such that the powers I, A and A^2 are linearly dependent, is a closed and lower-dimensional (semi-algebraic) subset of S.*

This theorem is a complement to McManus' (1992) result on the generic identifiability of nonlinear parametric models. For the social networks context, the key intuition for generic identifiability is that since A is assumed to be known *a priori*, this knowledge is the equivalent of a large number of coefficient restrictions on the coefficients in the reduced form representation of individual behaviors. These restrictions are rich enough that, outside of nongeneric cases, they permit identification of k, c, d and J.

iii. Unobserved component-specific fixed effects

The analog to group-level unobservables in the linear in means model in networks is component-level unobservables. If individual outcomes contain unobservables that are correlated among individuals belonging to the same component they may be treated as fixed effects in the stochastic structure of (47), producing

$$\omega_i = k + cx_i + d\sum_{j \neq i} a_{ij}x_j + J\sum_j a_{ij}\omega_j + \alpha_g + \varepsilon_i \tag{50}$$

with error structure

$$\mathrm{E}(\varepsilon_i | \alpha_g, (x_i)_{i \in g}, A) = 0, \tag{51}$$

where α_g is a component-specific fixed effect. This can be thought of as a model of interacting in groups, in which the groups themselves have internal social structure.

[33] A semi-algebraic set is a set which can be described as the solutions to a finite number of polynomial inequalities. The set of $n_V \times n_V$ matrices such that for all J, $(I - JA)^{-1}$ exists is a semi-algebraic set in $R^{n_V^2}$. Semi-algebraic functions are functions whose graphs are semi-algebraic sets. Every semi-algebraic set is the union of a finite number of disjoint open C^∞ manifolds. The dimension of a semi-algebraic set is the largest of the dimensions of these manifolds. For more on semi-algebraic geometry see Bochnak, Coste, and Roy (1998).

Little work has been done on this problem. We know of only the identification results of Bramoullé, Djebbari, and Fortin (2009). Since their model is linear, component-specific fixed effects can be differenced away.[34] In principle, this differencing can be done in many different ways, of which Bramoullé *et al.* discuss two. "Local differencing" subtracts from each individual's behavioral equation the average of those who directly influence him. "Global differencing" subtracts from each individual's behavioral equation the average of those in the connected component to which the individual belongs. A third differencing strategy not yet studied is to subtract from each individual's behavioral equation the average of those to whom he is indirectly connected. Differencing entails loss of information, and so conditions for identification are stronger. But here too identification is determined by the network topology. In particular, Bramoullé *et al.* prove the following theorem:

Theorem 6. Identification of social interactions in linear network models with component-specific fixed effects. *For the social interactions model described by equations* (46), (50) *and* (51), *assume that* $Jc + d \neq 0$. *With local differencing, a necessary and sufficient condition for identification of the parameters k, c, d and J is that the matrices I,* A, A^2 *and* A^3 *are linearly independent.*

Identification with global differencing is marginally simpler. Bramoullé *et al.* show that the independence condition of theorem 6 is sufficient but not necessary, and offer some weaker necessary and (distinct) sufficient conditions.

An unfortunate necessary assumption for their analysis is the conditioning in the expectation of the error term. We would expect models like (48) to be most interesting for those situations where network formation is endogenous, but the expectation assumption rules out any model wherein unobservable individual characteristics affect both network relations and individual behaviors.

iv. Self-selection in social network models

Investigating self-selection in social network models requires modelling the co-evolution of networks and behavior. Although the growth of networks has been studied empirically, and evidently behavior on networks is a well-established subject, the joint evolution of both has rarely been touched upon.[35] In particular, the econometric issues posed by endogenous network formation are briefly discussed by Jackson (2008, p. 437).[36]

[34] Bramoullé, Djebbari, and Fortin (2009) refer to these fixed-effects as "network-specific". "Component-specific" is a more precise description.

[35] These issues have come up in the study of online commmunities. See Crandall, Cosley, Huttenlocher, Kleinberg, and Suri (2008).

[36] See Bala and Goyal (2000) and Jackson and Wolinsky (1996) for notable contributions, and Jackson (2008, ch. 6) for an extensive treatment of several other works.

Weinberg (2008) proposes a model which is interesting for its treatment of a hierarchical and partially endogenous social structure. People belong to a group, whose aggregate exogenous characteristics and endogenous choices affect individual behavior, which is derived from a preference structure similar to that of appendix 1. But in contrast to modeling of group choice as discussed above in section 3.v and further pursued in section 5, Weinberg bypasses the treatment of groups as discrete objects and assumes instead that each group member can form a desired set of contextual effects by choosing to actively interact with (in his language, associate with) a subset of the group. In other words, he presumes that social structure arises from the direct choices of desirable contextual (and endogenous social) interactions by each individual rather than from the choice of group memberships from which social interactions are experienced. The objective of this modeling assumption is to allow for very general subgroups to form within predefined groups. For example, one can imagine distinct friendship networks forming within a classroom. This kind of model makes possible a rich set of identification strategies even though it is basically a linear in means model because the possibility of endogenous associations introduces nonlinear relationships into the behavioral equations. Weinberg identifies patterns of behavior that are not present when association is exogenous. For instance, Weinberg's model exhibits effects of group size not present in the linear in means model because larger groups offer more opportunities for reinforcement through association with likeminded individuals. The possibility of association introduces nonlinearities into the magnification effects of social interaction. For instance, Weinberg's model is consistent with the finding (Angrist and Lang, 2004) that relocating individuals at risk for certain behaviors or outcomes to less risky groups may have little impact because the at-risk individuals may associate differentially with those who engage in risky behavior.

For all of its virtues, Weinberg (2008) does not fit naturally into the frameworks that have emerged in the social networks literature. Subgroup selection is modeled crudely when compared to the body of social networks structures that have been studied. The modeling strategy carefully avoids the difficult issues of network topology and its effects on the diffusion of social action. Nonetheless, the development of panel data sets such as Add Health and those available from large online communities make spatial self-selection a topic of increasing importance.

v. Spatial econometrics specifications of social interactions

A close relationship exists between social interactions and spatial econometrics models. Equation (47) implies the classic Cliff-Ord spatial autoregressive (SAR) model with one spatial lag, as the special case of $d = 0$ where the dimension of endogenous outcomes is equal to the number of spatial units. These can be states, counties, parcels of land, etc. When instead of spatial units, the model refers to individuals, one has a social interactions model. The social interactions literature has recently sought to exploit the relationship. See

Lee (2007), who explored this link formally, and Lee, Liu, and Lin (2010).[37] In addition, the spatial econometrics literature has made important advances in terms of allowing for spatial autocorrelation in error structures: see Kapoor, Kelejian, and Prucha (2007) and Kelejian and Prucha (2010) for recent examples of advances in the study of spatial environments under weak error assumptions and Anselin (2010) for a review of the area. Spatial econometrics models have a long tradition in geography where the weights attached to different observations are motivated in terms of various distance concepts. For example, if the units of observations are counties, one may wish to account not only for adjacency but also for distance between their main population centers. Adding contextual effects, as in Lee (2007), brings the model closer to standard social interactions models.

An interesting predecessor to the new spatial econometric approaches to social interactions is Topa (2001). Topa's focus is on the exchange of information about labor market opportunities. He draws upon spatial epidemiological models in his study of job search. Unemployed individuals learn about job opportunities from, and employed individuals pass on job opportunities to, their immediate neighbors. Employed individuals become unemployed at an exogenously given rate, while unemployed individuals become employed at a rate proportional to the number of their employed neighbors. In its simplest form, this is a contact process, a network version of the basic SIR (susceptibles, infectives, and removables) model of disease spread.[38] While the model Topa estimates is somewhat different, it shares two important features: correlations across sites are positive, and they diminish geometrically with distance. Vertices in the network are census tracts in Chicago, and the dependent variable is the tract unemployment rate, which is modeled to move in a manner that is a metaphor for the aggregation of the contact process over individuals in each tract. The model is a first-order Markov chain. A state of the chain is an unemployment rate for each census tract. The transition probability is described by specifying a probability distribution for each tract i, conditioned on the current unemployment of tract i and its exogenous characteristics (education levels, ethnic composition and the like), and the current unemployment rate of its neighbors.

Topa does not have time series data, and so his estimation strategy is to fit the invariant distribution. Since the likelihood function is intractable, he estimates the model by indirect inference with an SAR auxiliary model. In short, he estimates the auxiliary model with the realized data. He then simulates the structural model for different parameter values, and estimates the auxiliary model with these simulated data sets. The indirect inference estimate is the set of parameter values that minimize a distance between the simulated-data and real-data parameter estimates of the auxiliary model. Point identification in the structural model is likely but difficult to establish

[37] Boucher, Bramoullé, Djebbari, and Fortin (2010) estimate the full Lee (2007) model with group-level unobservables using data on student achievement from Quebec secondary schools and find evidence of endogenous peer effects while also controlling for contextual effects and group unobservables in the form of fixed effects.

[38] See Daley and Gani (1999) for a good exposition of the SIR model.

since the likelihood function is not directly accessible. Observed spatial correlation may be due to endogenous spillovers, social interactions, or to spatial correlation in the neighborhood characteristics. Topa asserts that the nonlinear nature of the model makes identification likely. We have already shown cases where nonlinearity facilitates identification in our discussion of the nonlinear in means and hierarchical models, and will see other cases below, so the assertion seems plausible. If the model were linear, arguments like those of Bramoullé, Djebbari, and Fortin (2009) would apply. Assuming that the structural model is identified, the indirect inference estimator identifies the parameters if the minimum distance estimate is unique and if the map from structural parameters to parameters for the auxiliary model is one to one, and is locally onto the range of auxiliary parameter values in a neighborhood of the image of the true structural parameter value. The latter condition amounts to a rank requirement which Topa tests. Conley and Topa (2007) extend Topa (2001) by allowing for the use of individual-level information in the analysis of a variant of Topa's structural model, using a combination of calibration and estimation rather than indirect inference with an auxiliary equation. Together, Topa (2001) and Conley and Topa (2007) raise and begin to address the very real problem of how to assess identification when a structural model is not immediately accessible, an area which needs much work.

Lee, Liu, and Lin (2010) is a significant advance in the econometrics of social networks and spatial models. It generalizes Lee (2007) by allowing for group unobservables and correlated disturbances of connected individuals. Spatial autocorrelations in the error structure of their model are accounted for by assuming that the vector of shocks for a given component l (comprised of n_l members), ε_l, consists of the sum of group-specific fixed effects and stochastic components that satisfy

$$\varepsilon_l = \rho A_l^* \varepsilon_l + \nu_l,$$

where A^* is an exogenous and nonstochastic $n_l \times n_l$ nonnegative error-interactions matrix that need not coincide with A.[39] The parameter ρ is the spatial autocorrelation coefficient, and ν_l is a n_l-vector of i.i.d. individual-specific shocks. This error specification may be thought of as a generalization of a number of previous studies. Relative to Lee (2007), in which an individual in a group interacts with all other group members with equal weights (and identification is ensured by different group sizes), Lee, Liu, and Lin (2010) allows different individuals to have their own social groups, defined by the respective social interactions matrices A_l.

This model also extends the standard SAR model by allowing for contextual effects and group unobservables. These authors characterize the identification conditions of such an augmented SAR model based on features of the network structure, the role of exogenous variables, and the presence of correlated disturbances. Their results are

[39] In their analysis, nondiagonal entries are assumed to be symmetric and positive, but diagonal entries are 0.

broadly similar to those of Bramoullé *et al*. Lee, Liu and Lin also include asymptotic and estimation issues. In particular, they propose a method of eliminating fixed effects that may be used with a general adjacency matrix for which simple averaging might not work. They suggest a method which eliminates the fixed effects and allows the estimation of the remaining parameters of interest via a quasi maximum likelihood method that explores the row-normalization property of the social interactions matrices. Lin (forthcoming) estimates a version of the Lee *et al*. model using Add Health data with student academic achievement as the endogenous variable and finds strong evidence for both endogenous and contextual effects, even after controlling for school-grade fixed effects, and significant spatial autocorrelation in the disturbances. The estimation results differ greatly with or without school-grade fixed effects, which suggests that failure to correct for correlated unobservables may produce severe bias. Such differences may also be due to selection bias as well, in that individuals choose their peers. Systematic econometric methods of accounting for self-selection bias in social network formation are not yet available.

vi. From econometrics to applications

We end this section with two illustrations of how the types of models we have discussed have been applied in empirical work. Calvó-Armengol, Patacchini, and Zenou (2009) is a good illustration of how network methods have been employed. Using the Add Health data set, these authors estimate individual school performance as a function of the topology of their friendship networks, while controlling for individual characteristics. Individuals' friends always lie in the same school as themselves. They estimate a restricted variant of equation (47) in which $J = 0$. This model is generalized, however, by allowing the unobserved individual-level heterogeneity to also be related to the structure of the component level interactions. For each component l, the vector of errors ε_l for members of component l obeys

$$\varepsilon_l = \varrho A\imath + \rho A\varepsilon_l + \nu_l, \tag{52}$$

where A is the same adjacency matrix that links observable characteristics across individuals, ϱ measures the mean effect of the number of direct neighbors for each individual ("best friends" according to the Add Health questions), which is given by $A\imath$, ρ denotes a spatial autocorrelation coefficient in the ε's, and ν_l is again an n_l-vector of i.i.d. individual-specific shocks.[40] Since the error structure in (47), ε_l, represents individual outcomes that are not explained by individual characteristics x and contextual effects Ax, these authors reason that it proxies for peer interactions. Their estimation of the stochastic structure subsumes social interactions into the estimation of (ϱ, ρ) in

[40] $\imath$ is an $n_l \times 1$ vector of 1's. For the Add Health data set, A is an adjacency matrix where a 1 means that at least i or j has designated the other as a best friend, so $A\imath$ is the number of friends of each member of the component.

(52), the former because it controls for the number of best friends and the latter because it reflects how each individual's unobservable shock is affected by those of his friends. These authors interpret (ϱ, ρ) as expressing peer effects, which is not a standard use of terminology. From the perspective of the distinctions we have drawn between types of social interaction effects, ϱ and ρ parameterize the strengths of different contextual effects since neither refers to direct interdependences of choices per se.

Bayer, Ross, and Topa (2008) is a good example of how spatial reasoning can be used to identify social interactions. This paper develops a novel strategy to identify and estimate local referral interactions in the labor market. In order to identify labor market referrals from other spatially correlated interactions, the authors estimate the propensity to work together (in the same city block) for pairs of workers who reside in a given city block, controlling for the baseline propensity to work together for residents in nearby blocks (which the authors call a reference group of blocks).

The crucial identifying assumption here is that workers can choose residential locations down to a group of blocks, but do not purposefully choose among the individual blocks in the group because of block-specific characteristics. Therefore, conditional on sorting at the group of blocks level, the assignment of individuals to specific blocks is independent of group-specific characteristics. The authors use this conditional independence to identify local interactions with respect to labor market referrals. Specifically, let i and j denote individuals who reside in the same Census block group but do not belong to the same household. The outcome of interest is the binary variable ω_{ij} which indicates whether or not i and j work in the same Census block.[41] Further, δ_{ij}^{b} is a dummy variable that equals 1 if i and j reside in the same Census block, x_{ij} denotes a vector of socio-demographic characteristics for the pair i, j, and η_g denotes a reference group fixed effect which serves as the baseline probability of an employment match for individuals living in the same block group. The proposition that block-level interactions occurs in labor market referrals is defined via the regression

$$\omega_{ij} = \beta x_{ij} + (\alpha_0 + \alpha_1 x_{ij})\delta_{ij}^{b} + \eta_g + \varepsilon_{ij}. \tag{53}$$

The Bayer *et al.* test for the presence of social interactions due to proximity reduces to testing for the statistical significance of α_0 and α_1 in (53). The observable pair covariates term βx_{ij} controls for individual-specific reasons why two individuals work on the same block and η_g controls for any unobserved heterogeneity that occurs at the block group level and affects employment location. For example, η_g may be argued to control for features of the urban transportation network that might induce clustering in both residence and work location. The empirical strategy of Bayer *et al.* addresses several

[41] Census blocks are measured for the Boston area using 1990 US Census data. As Bayer *et al.* remark, Census blocks approximate physical city blocks. This definition creates some measurement questions since it means that two individuals who live opposite one another on given street are defined as living on different blocks.

additional potential pitfalls, including possible sorting below the block level and the possibility of reverse causation due to co-workers giving referrals on desirable residential areas. They find large block-level social interactions effects on employment location, especially among individuals who are socioeconomically similar.

These examples illustrate the need for a general method of testing whether a particular specification is compatible with the data in the general sense that the residuals do not contain evidence of omitted network structure. A study by de Graaff, Florax, Nijkamp, and Reggiani (2002) suggests a general specification test which could be applied to test whether the residuals contain "omitted structure" that could improve the predictive ability of estimated models. Their approach is rather abstract and is a test for the independence of the residuals of a particular specification, whereas what one ideally wants is a test for exchangeability of the residuals of a specification. It is still potentially very useful as a first step towards testing for exchangeability of the residuals. This is so because if the de Graaff *et al.* test of the null of independence is passed by the residuals, then this evidence is consistent with exchangeability of the residuals. Nevertheless, despite this caveat, the de Graaff *et al.* test is a potentially useful specification-testing tool in the area of social econometrics. Furthermore, we believe that, in general, the econometric methodology from spatial econometrics remains under exploited in social interactions econometrics. See Anselin (2010) for a very recent review of the subject.

vii. Social networks with unknown network structure[42]

All the results in this section so far have taken the social network matrix A as known. This severely restricts the domain of applicability of existing identification results on social networks. We finish this section by considering how identification may proceed when this matrix is unknown. In order to do this, we believe it is necessary to consider the full implications of the interpretation of linear social interactions models as simultaneous equations systems. While this interpretation is given in studies like Bramoullé *et al.*, the full implications of this equivalence have not been explored. This is evident if one observes that the matrix form of the general social networks model may be written as

$$(I - JA)\omega = (cI + dA)x + \varepsilon \tag{54}$$

where for expositional purposes, the constant term is ignored. From this vantage point, it is evident that social networks models are special cases of the general linear simultaneous equations system of the form

$$\Gamma\omega = Bx + \varepsilon. \tag{55}$$

Systems of this type, of course, are the focus of the classical identification in econometrics, epitomized in Fisher (1966) and comprehensively summarized in Hsiao (1983). One can go

[42] This section was inspired by comments by Gary Becker and especially James Heckman.

further and observe that the assumption that the same network weights apply to both contextual and endogenous social interactions is not well motivated by theory, and regard equation (55) as the general specification of a linear social networks model where the normalization $\Gamma_{ii} = 1$ for all i is imposed. From this vantage point it is evident that the distinction between J and A is of interest only when A is known *a priori*, as is the case both for the linear in means model and the more general social networks framework.

Following the classical literature, one can then think of the presence or absence of identification in terms of whether particular sets of restrictions on (55) produce identification. All previous results in this section are examples of this perspective but rely on the very strong assumption of a particular way of imposing these restrictions, i.e., $\Gamma = I - JA$ and $B = cI + dA$ for known A. Note that the results we have described do not employ information on the variance covariance matrix of the reduced form error structure, which is one source of identifying information and the basis for Graham's (2008) results. The simultaneous equations perspective makes clear that the existing results on identification in linear social networks models can be extended to much richer frameworks. We consider two classes of models in which we interpret all agents $i = 1,\ldots, n_V$ as arrayed on a circle. We do this so that agents 1 and n_V are immediate neighbors of one another, thereby allowing us to work with symmetric interaction structures.

First, assume that each agent only reacts to the average behaviors and characteristics of his two nearest neighbors, but is unaffected by anyone else. This is a linear variation of the model studied in Blume (1993). In terms of the matrices Γ and B, one way to model this is to assume that, preserving our earlier normalization, $\Gamma_{ii} = 1$ and $\Gamma_{ii-1} = \Gamma_{ii+1} = \gamma_1$ for all i, $\Gamma_{ij} = 0$ otherwise; $B_{ii} = b_0$, $B_{ii-1} = B_{ii+1} = b_1$ for all i, and $B_{ij} = 0$ otherwise, where here (and for the remainder of this discussion), all indices are mod n_V. The model is identified under theorem 5 since the nearest neighbor model may be interpreted via the original social networks model via restrictions on A. For our purposes, what is of interest is that identification will still hold if one relaxes the symmetry assumptions so that $\Gamma_{ii-1} = \gamma_{i-1}$, $\Gamma_{ii+1} = \gamma_{i1}$, $B_{ii} = b_{i0}$, $B_{ii-1} = b_{i-1}$ and $B_{ii+1} = b_{i1}$. If these coefficients are nonzero, then the matrices Γ and B fulfill the classical rank conditions for identification, cf. Hsiao (1983, theorem 3.3.1) and one does not need to invoke theorem 4 at all. Notice that it is not necessary for the interactions parameters to be the same across agents in different positions in the network. Relative to Bramoullé *et al.*, what this example indicates is that prior knowledge of A can take the form of the classical exclusion restrictions of simultaneous equations theory. From the vantage point of the classical theory, there is no need to impose equal coefficients across interactions as those authors do. Imposition of assumptions such as equal coefficients may be needed to account for aspects of the data, e.g., an absence of repeated observations of individuals. But if so, then the specification of the available data moments should be explicitly integrated into the identification analysis, something which has yet to be done. Further, data sets such as Add Health, which produce answers to binary questions concerning

friends, are best interpreted as providing 0 values for a general A matrix, but nothing more in terms of substantive information.

This example may be extended as follows. Suppose that one is not sure whether or not the social network structure involves connections between agents that are displaced by 2 on the circle, i.e., one wishes to relax the assumption that interactions between agents who are not nearest neighbors are 0. In other words, we modify the example so that for all i, $\Gamma_{ii} = 1$, $\Gamma_{ii-1} = \Gamma_{ii+1} = \gamma_1$, $\Gamma_{ii-2} = \gamma_{i-2}$, $\Gamma_{ii+2} = \gamma_{i2}$, $\Gamma_{ij} = 0$ otherwise, $B_{ii} = b_{i0}$, $B_{ii-1} = b_{i-1}$, $B_{ii+1} = b_{i1}$, $B_{ii-2} = b_{i-2}$, $B_{ii+2} = b_{i2}$, and $B_{ij} = 0$ otherwise. If the nearest neighbor coefficients are nonzero, then by Hsiao's theorem 3.3.1 the coefficients in this model are also identified regardless of the values of the coefficients that link non-nearest neighbors. This is an example in which aspects of the network structure are testable, so that relative to Bramoullé *et al.* one does need to exactly know A in advance in order to estimate social structure. The intuition is straightforward, the presence of overlapping network structures between nearest neighbors renders the system overidentified: so that the presence of some other forms of social network structure can be evaluated relative to it. This form of argument seems important as it suggests ways of uncovering social network structure when individual data are available, and again has yet to be explored. Of course, not all social network structures are identified for the same reason that without restrictions, the general linear simultaneous equations model is unidentified. What our argument here suggests is that there is much to do in terms of uncovering classes of identified social networks models that are more general than those that have so far been studied.

For a second example, we consider a variation of the model studied by Bramoullé *et al.*, which involves geometric weighting of all individuals according to their distance; as before we drop the constant term for expositional purposes. Specifically, we consider a social networks model

$$\omega_i = cx_i + d\sum_{j\neq i} a_{ij}(\gamma)x_j + J\sum_{j\neq i} a_{ij}(\gamma)\omega_j + \varepsilon_i.$$

The idea is that the weights assigned to the behaviors of others are functions of an underlying parameter γ. In vector form, the model is

$$\omega = cx + dA(\gamma)x + JA(\gamma)\omega + \varepsilon, \tag{56}$$

where

$$A(\gamma) = \begin{pmatrix} 0 & \gamma & \gamma^2 & \cdots & \gamma^k & \gamma^k & \gamma^{k-1} & \cdots & \gamma^2 & \gamma \\ \gamma & 0 & \gamma & & \cdots & \gamma^k & \gamma^k & \gamma^{k-1} & \cdots & \gamma^2 \\ & & & & \vdots & & & & & \\ \gamma & \gamma^2 & & \cdots & & & & & \gamma & 0 \end{pmatrix}. \tag{57}$$

Following Bramoullé *et al.*, x is a scalar characteristic. The parameter space for this model is $\mathcal{P} = \{(c, d, J, \gamma) \in \mathbf{R}^2 \times \mathbf{R}_+ \times [0, 1)\}$. The reduced form for this model is

$$\omega = (I - JA(\gamma))^{-1}(cI + dA(\gamma))x + (I - JA(\gamma))^{-1}\varepsilon.$$

Denote by $F : \mathcal{P} \to \mathbf{R}^{n_V^2}$ the map

$$F(c, d, J, \gamma) = (I - JA(\gamma))^{-1}(cI + dA(\gamma)). \tag{58}$$

The function F characterizes the mapping of structural model parameters (c, d, J, γ) to reduced form parameters. We will establish what Fisher (1959) calls complete identifiability of the structural parameters from the regression coefficients for the reduced form. By this he means that each reduced form parameter vector is derived from only a finite number of structural parameter vectors, i.e., that the map from structural models to reduced form models is finite-to-one.

The behavioral model described by (56) is nonlinear in the parameters because d, γ and J interact multiplicatively. This is nonetheless a natural model, as it is the simplest way to discount individual effects by distance. The following complete identification result holds for this model:

Theorem 7. Identification of the linear social networks model with weights exponentially declining in distance. *Suppose that the number of individuals n_V is at least 4. Then for all $(c, d, J, \gamma) \in \mathcal{P}$,*

i. *if $I - JA(\gamma)$ is nonsingular, $c + d \neq 0$ and $\gamma \neq 0$, then the cardinality of $F^{-1}(F(c, d, J, \gamma))$ is no more than $2(n_V - 1)$.*

ii. *The events $J = d = 0$ and $\gamma = 0$ are observationally equivalent. In this case, $F(c, d, J, \gamma) = cI$.*

Part i. of theorem 7 says the following: Each structural parameter vector is observationally equivalent to at most $2n_V - 3$ other structural parameter vectors in the sense that they all generate the same reduced form. As such, while point identification may not be achieved, any true structural parameter vector fails to be identified relative to at most $2n_V - 3$ alternatives. Notice that complete identification is stronger than local identification. Local identification implies that for the true structural parameters, there exists an open neighborhood of these parameters that does not contain any observationally equivalent structural parameters. The set of observationally equivalent structural parameters could nonetheless be countable. Complete identification requires that the set be finite, which implies local identification. Part ii. notes that if there are no social interactions, this imposes sufficiently strong restrictions on the reduced form parameters to identify both c and also requires that the matrix of reduced form parameters is proportional to an identity matrix. We believe that refinements of theorem 7.i are possible and leave this to future work.

These examples illustrate how the results of sections 3 and 4 are far from exhaustive in understanding the identification of linear social interactions models.

5. DISCRETE CHOICE MODELS OF SOCIAL INTERACTIONS

In this section, we consider identification for discrete choice models. Identification conditions for discrete choice models will prove to be conceptually quite different than the conditions that apply to linear models. Some reasons are trivial. For example, discrete choice models, because they involve probabilities, are inherently nonlinear and as we have discussed, nonlinear models have very different identification conditions than linear ones. Other differences will prove to be more subtle.

i. Binary choice: basic structure

We first focus on binary choice models of social interactions. These have been the primary focus of theoretical work. Early theoretical studies include Blume (1993), Brock (1993), Durlauf (1993) and Glaeser, Sacerdote, and Scheinkman (1996). Recent contributions which generalize these earlier analyses in terms of the timing and network structure of interactions as well as in terms of belief formation include Bisin, Horst, and Özgür (2006), Horst and Scheinkman (2006) and Ioannides (2006).

Durlauf and Ioannides (2010) provide an overview of some of the interesting theoretical properties of these models. We provide a brief summary in appendix 3. Unlike the linear in means model, discrete choice models of social interactions can exhibit multiple equilibria. Multiple equilibria are a well-known challenge to identification in contexts such as industrial organization but will turn out to facilitate identification of social interactions.[43] A second important property of discrete choice models with social interactions is that they may exhibit bifurcations around certain parameter values, which essentially means that there will exist threshold values for model parameters such that if the parameters take values below the threshold, the environment under study will exhibit one set of qualitative properties whereas if the parameters take on values greater than the threshold, the environment will exhibit qualitatively different properties. Since these properties are observable, identification can depend on the values of the model parameters. Appendix 3 illustrates these properties under the case of logistic errors, using a model due to Brock and Durlauf (2001a).

Identification for binary choice models has been studied in detail by Brock and Durlauf (2001a,b, 2007); other contributions include Soetevent and Kooreman (2007). We follow Brock and Durlauf (2001a,b) for the development of an initial structure and indicate how subsequent analyses have relaxed assumptions relating to their formulation. Choices are coded so that they belong to the set $\{-1, 1\}$. If the context is teenage pregnancy, then $+1$ can denote *had a child while a teenager* while -1 can

[43] See Berry and Tamer (2006) for an overview, and Beresteanu, Molchanov, and Molinari (2008) and Galichon and Henry (2008) for recent advances in partial identification.

denote *did not have a child while a teenager*. To interpret these choices as the outcomes of a decision problem, we define individual-specific payoff functions $V_i(\omega_i)$.

An econometrically implementable choice structure is implemented by assuming the difference between the payoffs for the two choices is additive in the different factors that have been defined for the linear model, i.e.,[44]

$$V_i(1) - V_i(-1) = k + cx_i + dy_g + Jm_{ig}^e - \varepsilon_i. \tag{59}$$

Note that unlike the linear in means model, it is not necessary to require $J \in [0, 1)$ because here it has a different interpretation, as a utility parameter. We will almost always discuss the model as if $J > 0$ as this is the standard case of interest in the literature, but theory imposes no natural upper bound on J. Analogous to our initial analysis of the linear in means model, we make two error assumptions. First, the expected value of the unobservable ε_i term is independent of observable features of the individual and any features of his group:

$$F(\varepsilon_i|(x_j)_{j\in g}, y_g, i \in g) = F_\varepsilon(\varepsilon_i). \tag{60}$$

Second, any pair i and j of the errors are conditionally independent within and across groups:

$$F(\varepsilon_i, \varepsilon_j|(x_k)_{k\in g}, y_g, i \in g, (x_l)_{l\in h}, y_h, j \in h) = F_\varepsilon(\varepsilon_i)\cdot F_\varepsilon(\varepsilon_j) \quad \text{unless } i = j \text{ and } g = h. \tag{61}$$

These conditions are the analogs of the error restrictions (8) and (9) that were initially imposed on the linear in means model. These conditions are substantially stronger than those that appear in the linear in means model as they impose conditional independence rather than set certain conditional expectations equal to 0. They are also stronger than needed for identification proofs. It is well understood in the discrete choice literature that median restrictions can play a role analogous to expected value restrictions in linear models.[45] We make them here for ease of exposition and to link directly to theoretical results as developed in Brock and Durlauf (2001a,b, 2007).

The decision problem for this binary choice context is simple: individual i chooses $+1$ if and only if $V_i(1) - V_i(-1) \geq 0$. Hence

$$\begin{aligned}\mu(\omega_i = 1|x_i, y_g, i \in g) &= \mu(V_i(1) - V_i(-1) \geq 0) \\ &= \mu(\varepsilon_i \leq k + cx_i + dy_g + Jm_{ig}^e) = F_\varepsilon(k + cx_i + dy_g + Jm_{ig}^e).\end{aligned}$$

[44] The payoff differential is written in terms of $-\varepsilon_i$ for algebraic convenience. See the derivation of choice probabilities in appendix 3.

[45] See Horowitz (2009) for an extended treatment.

As before, the model is closed by imposing an equilibrium condition on beliefs. Each person is assumed to know γ_g, F_ε, and $F_{x|g}$, the empirical within-group distribution of x_i. When the population size is large, equilibrium requires that the expected value of the average choice level in the population, given this information, is defined by

$$m_g = 2\int F_\varepsilon(k + cx + d\gamma_g + Jm_g)dF_{x|g} - 1. \tag{62}$$

There typically does not exist, a closed form solution for m_g.

McFadden (1974) observed that the logit, probit, and similar discrete-choice models have two interpretations. The first interpretation is that of individual random utility. A decisionmaker draws a utility function at random to evaluate a choice situation. The distribution of choices then reflects the distribution of utility, which is the object of econometric investigation. The second interpretation is that of a population of decisionmakers. Each individual in the population has a deterministic utility function. The distribution of choices in the population reflects the population distribution of preferences. Brock and Durlauf (2001a) (and theoretical models such as Blume (1993)) extend this idea to games. One interpretation of this game theoretic approach is that the econometrician confronts a population of random-utility maximizers whose decisions are coupled. These models extend the notion of Nash equilibrium to random-utility choice. The other interpretation views an individual's shock as known to the individual but not to others in the population (or to the econom trician). In this interpretation, the Brock-Durlauf model is a Bayes-Nash equilibrium of a game with independent types, where the type of individual i is the pair (x_i, ε_i). Information is such that the first component of each player i's type is common knowledge, while the second is known only to player i.

ii. Identification

Identification of the parameters in the binary choice model holds for very different conditions than were seen in the linear in means case. These differences derive from the nonlinear nature of the binary choice and do not require that the functional form F_ε is known *a priori*. The following theorem is proved in Brock and Durlauf (2007). We emphasize that the theorem's conditions are sufficient, not necessary, and were chosen to render the sources for identification transparent.

Theorem 8. Identification of the binary choice model with social interactions. *Suppose for the binary choice model social interactions described by equations* (59) *through* (62),

- **i**. *conditional on* (x_i, γ_g), *the random payoff terms* ε_i *are distributed according to* F_ε, *and* $F_\varepsilon(0) = 0.5$;
- **ii**. F_ε *is absolutely continuous with associated density* dF_ε. dF_ε *is positive almost everywhere on its support, the interval* (L, U), *which may be* $(-\infty, \infty)$;

iii. *for at least one group g, conditional on γ_g, each element of the vector x_i varies continuously over all* $\mathbf{R}$ *and* $\text{supp}(x_i)$ *is not contained in a proper linear subspace of* $\mathbf{R}^{\mathbf{R}}$;

iv. *γ_g does not include a constant; each element of γ_g varies continuously over all* $\mathbf{R}$*; at least one element of d is nonzero; and* $\text{supp}(\gamma_g)$ *is not contained in a proper linear subspace of* $\mathbf{R}^{\mathbf{S}}$.

Then k, c, d, J and F_ε are identified up to scale.

The intuition for why identification holds is as follows. Within a given group, $d\gamma_g + Jm_g$ is constant for all agents. Assumptions i)–iii) are sufficient to ensure that within that group, the parameter vector c and density function F_ε are identified up to scale. Identification of these objects up to scale was originally established by Manski (1988). The assumptions stated here allow for a proof structure that mimics Horowitz (2009). Assumption iv) ensures that k, d and J are identified up to scale. Identification of k is trivial if the other parameters are identified. The reason why d and J are identified is that the unbounded support on the γ_g element with a nonzero coefficient ensures that m_g and γ_g cannot be linearly dependent. This follows simply from the fact that m_g is bounded between -1 and 1. This bound is not driven by any functional form assumption but follows from the fact that the expected choice values are functions of the choice probabilities that are bounded between 0 and 1. Hence, the argument for identification is analogous to one of the basic reasons why bounds can be established on probabilities in the partial identification literature. (See Manski (2003) for a synthesis.) Note as well that this is not an identification at infinity argument.

This theorem extends Brock and Durlauf (2001a,b) who proved identification when F_ε is a negative exponential distribution of the type used in appendix 3 and Brock and Durlauf (2006) who proved identification for general F_ε when F_ε is known *a priori*. Clearly, the conditions of this theorem can be relaxed. For example, if condition iii) holds for all groups, then one can allow for multiple $F_{\varepsilon g}$'s, i.e., different group-specific distributions. Similarly, one does not need unbounded supports for all regressors, rather what one needs is a large enough support for a nonlinear relationship between m_g and γ_g to ensure identification.

iii. Observability of actions

The identification results in Brock and Durlauf (2007) are sensitive to the assumption that individuals react to expected rather than realized behaviors of others. This follows from the assumption that an individual's random shock is observed only by himself. Soetevent and Kooreman (2007) build a game theoretic model with a different assumption. They assume that each individual knows the other individuals' shocks, that shocks are invisible only to the econometrician. Thus in equilibrium, each individual's expectation of the average choice of others will be the realized average choice of others. Soetevent and Kooreman have replaced the incomplete information and Bayes-Nash equilibrium of Brock and Durlauf (2001a) with complete information and Nash equilibrium. They justify their informational assumption

by presuming to study interactions in relatively small groups of given sizes in which choices of other individuals are assumed to be fully observed, and therefore an individual's payoff depends on the actual choice of others in his group, as opposed to expected choices as in Brock and Durlauf. This difference in information structure and the resulting equilibrium concept makes for an interesting contrast between the identification conditions in the two models. The equilibrium payoff difference equation (59) now becomes

$$V_i(1) - V_i(-1) = k + cx_i + dy_g + \frac{J}{n_g - 1}\sum_{j \neq i}\omega_{jg} - \varepsilon_i. \tag{63}$$

Like Brock and Durlauf, Soetevent and Kooreman focus on pure Nash equilibria with binary outcomes and estimate the model in effect as a system of simultaneous equations by means of simulation methods. Each individual choice is determined by the rule

$$\omega_{ig} = \begin{cases} 1 \text{ if } k + cx_i + dy_g + \dfrac{1}{n_g - 1}\sum_{j \neq i}\omega_{ig} > \varepsilon_i; \\ -1 \text{ if } k + cx_i + dy_g + \dfrac{J}{n_g - 1}\sum_{j \neq i}\omega_{jg} \leq \varepsilon_i. \end{cases}$$

Unlike the Brock and Durlauf model, for given values of parameters (k, c, d, J) and data (x_i, y_g), the decision rules for the individual agents may not produce unique strategy profiles.[46] This creates a very different multiple equilibrium problem than occurs in Brock and Durlauf, since in the latter, each agent has a unique strategy profile given the expected average choice m_g. Consequently, the normal approach of forming the likelihood function would not be appropriate in their case even when the ε_i's are independent.

Soetevent and Kooreman employ simulation-based estimation methods to compute the likelihood that any choice pattern would be observed. Their approach accounts for the potential multiplicity of noncooperative equilibria. For parameter values that generate multiple equilibria they assume that the equilibria are equally likely, which in turn guarantees statistical coherency of the model.[47] Simulation of the model over different regions of the parameter space allows for calculation of the number of equilibria for

[46] Soetevent and Kooreman (2007, pp. 602–3). This finding verifies the claim by Krauth (2006b) that with small (finite-size) social groups, the Brock and Durlauf model can exhibit multiplicity of strategy profiles whenever observed group behavior exerts any influence. The range of equilibrium group behavior depends on the size of the social group as well as its strength of influence.

[47] The assumption that all equilibria are equally likely is questionable. Blume and Durlauf (2003) show that in dynamic analogs of the Brock and Durlauf model, the percentage of time spent in the vicinity of the highest average utility equilibrium exceeds that of other equilibria; similarly Brock and Durlauf (2001b), for a version of the discrete choice model of social interactions in which the conditional probabilities of each choice depend on the realized choices of others, show that the equilibrium choice configuration will assign almost all probability to the social optimum as the population becomes large. While these analyses employ different microfoundations from Soetevent and Kooreman, they suggest that not all equilibria are equally likely. We thank James Heckman for discussion on this general issue.

draws of the ε_i's which are assumed to be i.i.d. normal. The procedure skirts the issue of exact identification of the model (no proof of identification is given) but provides a practical approach for implementation of their theoretical model. Also, in their actual estimation, Kooreman and Soetevent exclude contextual effects by setting $d = 0$.

The multiplicity of equilibria in Soetevent and Kooreman is very similar to types of multiplicity that have been studied in the industrial organization literature. Tamer (2003) launched a now thriving literature on multiple equilibria and partial identification by means of bounds for industrial organization contexts. This body of work, surprisingly, has had little contact with the social interactions literature. Clearly, both literatures would benefit from integration.

iv. Unobserved group effects

Unobserved group effects may be introduced in a fashion directly analogous to the linear in means model. Specifically, payoff differentials are described by

$$V_i(1) - V_i(-1) = k + cx_i + dy_g + Jm_g + \alpha_g - \varepsilon_i.$$

Here α_g is a fixed effect and equilibrium is required. Recall that individual agents are assumed to observe α_g while the analyst does not.

Without any restrictions on this fixed effect, it is evident that identification breaks down. Note that the presence of a fixed effect does not affect identification (up to scale) of c and F_ε. This holds because α_g is constant within a group and so is subsumed in the constant term. To see why the other parameters are not identified, observe that parameter values k, d and J are observationally equivalent to $\bar{k}$, $\bar{d}$ and $\bar{J}$, that is, for all $y_g \in \text{supp } y$,

$$k + dy_g + Jm_g + \alpha_g = \bar{k} + \bar{d}y_g + \bar{J}m_g + \bar{\alpha}_g$$

if one chooses $\bar{\alpha} = \alpha + Jm_g$ and $\bar{J} = 0$. Thus J and d are not identified. (See Brock and Durlauf (2007) for an elaboration.) We can therefore state:

Theorem 9. Nonidentification with unobserved group effects. *In the presence of unobserved group interactions whose properties are unrestricted, the parameters of the binary choice model with social interactions are not identified up to scale.*

In response to unobserved group effects, instrumental variables and differencing strategies are available just as occurs for linear models. Our remarks on instrumental variables for the linear in means model apply for the binary choice context as well and so are not repeated. Instead, we focus on two strategies, one which parallels the linear in means model and one which is new and only applies in the binary choice context.

a. Panel data

We first consider how panel data can be used to eliminate unobserved group effects for the binary choice model. Panel data, of course, allows one to consider differencing methods. The notion of differencing in panels for binary choice data is more subtle

than was the case for the linear in means model since it involves considering differences in probabilities across time. Chamberlain (1984) provides the generalization of differencing to discrete choice contexts. Identification of social interactions with differencing is studied in Brock and Durlauf (2007) who consider

$$V_{it}(1) - V_{it}(-1) = k + cx_{it} + dy_g + ey_{gt} + Jm_{gt} + \alpha_g - \varepsilon_{it}. \tag{64}$$

The vector y_{gt} is introduced in order to distinguish between those contextual effects that are time varying and those that are not. Applying Chamberlain's ideas on quasi-differencing of discrete data to models with social interactions, Brock and Durlauf verify a corollary to theorem 8.

Corollary 2. Identification of a subset of parameters with panel of the binary choice models of social interactions with fixed effects. *For the binary choice model with social interactions described by equations* (59)–(62), *assume that within period-choices are described by equation* (64) *and that the model equilibrium conditions hold period by period. If the assumptions of theorem 8 hold for all t, then c, e and J are identified up to scale whereas k and d are not identified.*

b. Partial identification

For binary choice models, Brock and Durlauf (2007) have proposed partial identification approaches to social interactions which involve weak assumptions on unobservables. We consider two examples. The partial identification arguments we develop are qualitatively different from those that typically appear in the econometrics literature. The reason is that we do not establish probability bounds. Rather, we show how certain empirical observations represent evidence of social interactions, even though parameter magnitudes cannot be bounded. The approach we describe is theory-dependent in the sense that it involves asking how the introduction of unobserved heterogeneity into various models affects their properties. Put differently, we are concerned with uncovering "footprints" of social interactions in heterogeneity-filled environments using various theoretical models as the basis for the analysis.

Our first example of a weak assumption is first order stochastic monotonicity of group level unobservables. We assume that y_g is measured so that $d \geq 0$. We denote the conditional distribution of the unobservable given y_g as $F_{\alpha_g|y_g}$. Letting $\succ$ denote first order stochastic dominance, and using $>$ when comparing vectors to mean that each element of one vector is greater than the corresponding element of the other, we assume

$$\text{if } y_g > y_{g'}, \text{ then } F_{\alpha_g|y_g} \succ F_{\alpha_{g'}|y_{g'}}. \tag{65}$$

This assumption is sufficient to produce partial identification of social interactions.

Theorem 10. Pattern reversals and partial identification of endogenous social interactions. *For the binary choice model with social interactions described by equations* (59)–(62) *suppose that the distribution of fixed effects exhibits first order stochastic dominance*

with respect to the contextual effects as characterized by equation (65). *If assumptions i)–iv) in theorem 8 hold and*

$$\gamma_g > \gamma_{g'} \text{ and } \mathrm{E}(m_g|\gamma_g) < \mathrm{E}(m_{g'}|\gamma_{g'}), \tag{66}$$

then it must be the case that $J > 0$ *and* J *is large enough to produce multiple equilibria.*

The term "pattern reversals" refers to the case where the observed characteristics of two groups suggest one ordering in their expected average outcomes, while the opposite ordering in fact holds.[48] This reversal of outcomes with respect to fundamentals can occur for three reasons. One possibility is that the observed outcome ordering is due to sampling error. This is irrelevant to identification because of the analogy principle. The second reason is that the unobserved group effects reverse the ordering that is implied by $d\gamma_g$. This is ruled out, in an expectations sense, by the stochastic dominance assumption. The only remaining reason for the pattern reversal is that there are multiple equilibria associated with m_g such that the low γ_g group has coordinated on the high-expected average outcome equilibrium whereas the high γ_g group has not. This is why the theorem requires multiple equilibria. To be clear, endogenous social interactions may be present when no pattern reversal occurs. All that can be said is that a pattern reversal in the presence of stochastic dominance in the sense of (66) is evidence of social interactions.

Our second example involves restricting the conditional density of the unobserved group interactions given observed group characteristics, that is, $dF_{\alpha_g|\gamma_g}$, via unimodality,

$$\text{for all } \gamma_g,\ dF_{\alpha_g|\gamma_g} \text{is unimodal.} \tag{67}$$

This assumption is sufficient to verify

Theorem 11. Partial identification of endogenous social interactions when the density of unobservables is unimodal. *For the binary choice model with social interactions described by equations* (59)–(62), *suppose that fixed effects are added as characterized by equation* (67). *If assumptions i)–iv) in theorem 8 hold, then*

i. *if* $J = 0$, *then* $dF_{m_g|\gamma_g}$ *is unimodal;*

ii. *if* $J > 0$ *is large enough to produce multiple equilibria for the binary choice model with social interactions, then* $dF_{m_g|\gamma_g}$ *is multimodal.*

This result also is based on multiple equilibria. In this case, the multiple equilibria produce the multimodality described in the theorem. Two observations should be made about this result. First, no analogous result exists for the unconditional density of expected outcomes, dF_{m_g}. The reason is that integrating $dF_{m_g|\gamma_g}$ over γ_g to produce dF_{m_g} would not necessarily preserve multimodality if it is present in the conditional density and, in

[48] Of course, expected group values are not directly observed. Our identification analysis replaces sample means with population means, following the analogy principle.

contrast, may spuriously produce it when it is absent from the conditional density. This follows from the nonlinear relationship between m_g and γ_g. Second, multimodality is sufficient but not necessary for multiple equilibria in $dF_{m_g|\gamma_g}$ as mixture densities are not necessarily multimodal.[49]

One message of these two examples is that unlike other contexts, multiple equilibria can facilitate the identification of social interactions. The reason for this is that endogenous social interactions are the only mechanism by which multiplicity can be generated in these environments.

These theorems illustrate a general feature concerning multiple equilibria and social interactions, namely that multiple equilibria do not represent an impediment to the identification of social interactions but rather may be of use to an analyst. This possibility was first recognized for the binary choice case in Brock and Durlauf (2001b). The theorems we have presented by no means exhaust the potential use of multiple equilibria to uncover social interactions. de Paula and Tang (2010), for example, argue that multiple equilibria, because they imply that observed choice probabilities are mixtures of equilibrium-specific choice probabilities, may be used to uncover the sign of endogenous interactions J. We discuss their work below.

v. Self-selection

Self-selection for discrete choice models has generally been handled using instrumental variables methods. The concerns we articulate about this strategy for the linear in means model apply to the discrete choice context as well. In parallel to the case of group level unobservables, Brock and Durlauf (2007) provide a number of partial identification results that hold under relatively modest assumptions.

To do this, Brock and Durlauf (2007) treat the membership question as the outcome of a matching problem and place some restrictions on the equilibria that emerge from the matching. Matching is assumed to occur with respect to an individual index A_i and a group index T_g, defined as

$$A_i = cx_i - \varepsilon_i \tag{68}$$

$$T_g = d\gamma_g. \tag{69}$$

In the context of peer effects in classrooms, A_i may be thought of as student ability and T_g as teacher quality. For simplicity, the individual characteristics x_i are assumed to be measured so that $c \geq 0$.

Individuals and groups are matched in the sense that higher group quality is associated with higher individual quality. With respect to the equilibrium matching process, Brock and Durlauf assume

[49] See Lindsay (1995, p. 4–5) for a nice example of a unimodel two-part mixture.

$$\text{For any pair of groups } g \text{ and } g', T_g > T'_g \Rightarrow F_{A|T_g} \succ F_{A|T_{g'}}. \tag{70}$$

This assumption is weaker than one which imposes strict assortative matching between better groups and higher ability individuals; the latter is predicted by models such as Becker (1973). The assumption is qualitatively consistent with a range of payoff functions that relate groups and individuals, see Sattinger (1993) for a survey of equilibrium matching problems. Note that (70) places an implicit restriction on $F_{\varepsilon|\gamma_g i \in g}$. This assumption on matching leads to theorem 12.

Theorem 12. Partial identification of endogenous social interactions under assortative matching. *For the binary choice model of social interactions* (59)–(62), *assume assortative matching as described by* (70). *Then* $\mathrm{E}(m_g|T_g) > \mathrm{E}(m_{g'}|T_{g'})$.

This theorem is useful as it indicates how the presence of endogenous social interactions may be inferred if $T_g > T_{g'}$ yet $\mathrm{E}(m_g|T_g) < \mathrm{E}(m_{g'}|T_{g'})$. This can only occur, under the specification we have assumed, if group g has coordinated on an equilibrium expected average choice level other than the largest of the possible equilibria associated with it, while group g' has coordinated on an equilibrium other than the lowest possible expected average choice level among those it could have attained. The existence of multiple equilibria immediately implies $J > 0$.

The use of assortative matching to facilitate identification may be extended to panel data. To do this, modify (68) and (69) so that $A_{it} = k + cx_{it} + \varepsilon_{it}$ and $T_{gt} = d\gamma_g + e\gamma_{gt}$ and that (70) holds period by period. Brock and Durlauf show:

Corollary 3. Equality of average outcomes with equal observable contextual effects. *Assume that the binary choice model of social interactions* (59)–(62) *holds for all t with equilibrium at each date and assortative matching as described by* (70). *If* $J = 0$ *or* $J > 0$ *but is sufficiently small that* m_{gt} *is unique, then* $\gamma_{gt} = \gamma_{gt'}$ *implies* $m_{gt} = m_{gt'}$.

vi. Beyond the binary choice model

a. Multinomial choice

Little econometric work has been done on multinomial choice models with social interactions; as far as we know, the only contributions are Brock and Durlauf (2002, 2006) and Bayer and Timmins (2007). Nevertheless these models seem important in many contexts. We develop the analog to the binary choice model and establish identification. Multinomial choice models with social interactions can exhibit multiple equilibria and bifurcations in parallel to those found in binary choice models. Appendix 3 provides a brief discussion.

To formulate the model, we consider an environment in which each member of a common group makes a choice l from a common choice set with L discrete possibilities, i.e., $\Omega_{ig} = \{0,\ldots, L-1\}$. The same choices are assumed to be available regardless of group. The common choice set assumption is without loss of generality, since if

agents face different choice sets, one can always assume their union is the common set and then specify that certain choices have payoff of $-\infty$ for certain agents. Individual utility is defined as

$$V_{ig}(l) = k_l + cx_i + d_l\gamma_g + Jp^e_{igl} + \beta^{-1}\varepsilon_{il}. \tag{71}$$

Here p^e_{igl} denotes agent i's expected value for the fraction of group g that chooses l. This generalizes the preference structure of the binary choice model to account for any number of choices. As before, β indexes the degree of heterogeneity in the random payoff term ε_{il}. We assume that these unobserved utility terms are independent and identically distributed with a common distribution function F_ε. In parallel to the binary choice case

$$F(\varepsilon_{il}|(x_m)_{m\in g}, \gamma_g, i \in g) = F_\varepsilon(\varepsilon_{il}),$$

and

$$\begin{aligned}&\text{for all } i, j, g, h, k, l \text{ such that not all of } i = j, g = h, k = l \text{ hold}\\ &F(\varepsilon_{ik}\varepsilon_{jl}|(x_m)_{m\in g}, i \in g, \gamma_g, (x_n)_{n\in h}, \gamma_h, j \in h) = F_\varepsilon(\varepsilon_{ik})\cdot F_\varepsilon(\varepsilon_{jl}).\end{aligned}$$

For this model, the probability that agent i makes a particular choice l is the probability that l produces the maximum payoff among all choices according to (71). This amounts to the joint probability defined by

$$\mu\begin{pmatrix} \varepsilon_{i0} - \varepsilon_{il} \le \beta(k_l + c_l x_i + d_l\gamma_g + J_l p^e_{igl} \\ \qquad\qquad - k_0 - c_0 x_i - d_0\gamma_g - J_0 p^e_{ig0}) \\ \qquad\vdots \\ \varepsilon_{iL-1} - \varepsilon_{il} \le \beta(k_l + c_l x_i + d_l\gamma_g + J_l p^e_{igl} \\ \qquad\qquad - k_{L-1} - c_{L-1} x_i - d_{L-1}\gamma_g - J_{L-1} p^e_{igL-1}) \end{pmatrix}.$$

Following an order-statistics argument[50] the probability of choosing l conditional on a particular realization of ε_{il} is

$$\prod_{j\neq l} F_\varepsilon\Big(\beta(k_l + c_l x_i + d_l\gamma_g + J_l p^e_{igl}) - \beta(k_j + c_j x_i + d_j\gamma_g + J_j p^e_{igj}) + \varepsilon_{il}\Big),$$

which immediately implies that the unconditional probability of the choice l is

$$p_{igl} = \int \prod_{j\neq l} F_\varepsilon\Big(\beta(k_l + c_l x_i + d_l\gamma_g + J_l p^e_{igl}) - \beta(k_j + c_j x_i + d_j\gamma_g + J_j p^e_{igj}) + \varepsilon\Big) dF_\varepsilon. \tag{72}$$

[50] Anderson, de Palma, and Thisse (1992, p. 36) provides a clean exposition.

In equilibrium, the aggregate choice probabilities of this general multinomial choice model are the solutions to

$$p_{gl} = \int\int \prod_{j \neq l} F_\varepsilon \Big(\beta(k_l + c_l x_i + d_l \gamma_g + J_l p^e_{igl}) - \beta(k_j + c_j x_i + d_j \gamma_g + J_j p^e_{igj}) + \varepsilon \Big) dF_\varepsilon dF_{x|g}. \tag{73}$$

Brock and Durlauf (2006) prove a general identification theorem for the multinomial choice model.

Theorem 13. Parametric identification for the multinomial choice model. *Let the true data generating process be given by* (71)–(73) *and assume that* F_ε *is known. Under the normalization* $k_0 = 0$, $c_0 = 0$, $d_0 = 0$, *and* $J_0 = 0$, *if*

i. *the mapping defined by equation* (73) *is globally one-to-one*,
ii. *the joint support of* x_i, γ_g *is not contained in a proper linear subspace of* $\mathbf{R}^{R+S}$,
iii. *the support of* γ_g *is not contained in a proper linear subspace of* $\mathbf{R}^S$,
iv. *no linear combination of elements of* x_i and γ_g *is constant*,
v. *for each individual i, conditional on* γ_g, x_i *is not contained in a proper linear subspace of* $\mathbf{R}^R$,
vi. *none of the elements of* γ_g *has bounded support*,
vii. *for all l,* p_{gl} *is not independent of g*,

then the vector of model parameters $(k_1, c_1, d_1, J_1, \ldots, k_{L-1}, c_{L-1}, d_{L-1}, J_{L-1})$ *is identified up to scale.*

We emphasize that these conditions are sufficient, not necessary. The careful reader will observe that these conditions have not been reconciled with the conditions for the binary choice model. We do not do this in order to avoid uninteresting modifications of the proofs of the original theorem.

Relative to the binary choice identification theorem, conclusion of the theorem is weaker as it requires that the error distribution is known *a priori*. As far as we are aware, there does not exist any multinomial generalizations of Manski (1988) so that the error distribution may be identified (up to scale) via intragroup variation in individual characteristics and associated behaviors.

Bayer and Timmins (2007) study a variation of the multinomial choice problem, which focuses on choices across locations. They thus consider a population that forms a single group. We omit the group index in describing their model. In terms of the error structure, they set $\beta = 1$ in (71) and assume that the error terms are double exponentially distributed,

$$\mu(\varepsilon_{il} \leq \zeta) = \exp(-\exp(-\beta\zeta + \gamma)). \tag{74}$$

In terms of preferences, they follow the industrial organization literature in allowing for coefficient heterogeneity; their implementation of this heterogeneity is the opposite of the formulation one finds for hierarchical models in that the heterogeneity is

determined by individual characteristics. In addition, they allow for unobserved choice-specific fixed effects. This produces choice-specific payoffs

$$V_i(l) = d_{il}z_l + J_{il}p_{il} + \xi_l + \varepsilon_{il},$$

where

$$d_{il} = d + D_l x_i, \quad J_{il} = J + J_l x_i,$$

and ξ_l is an unobserved location-specific effect. Bayer and Timmins use the functional form assumption (74) to construct instruments for estimation of this model. Their approach is a variant of models that all fall under the approach pioneered by Berry, Levinson, and Pakes (1995). An interesting aspect of Bayer and Timmins' work is that they focus on identification power that derives from changes in substitution patterns in multinomial choice models.

b. Duration models

A number of studies have sought evidence in dynamic contexts based on duration and optimal stopping problems. Brock and Durlauf (2001b) first discussed this approach to modeling social interactions, albeit briefly. Sirakaya (2006) studies recidivism under the assumption the individual hazard function for an individual probationer depends on individual and neighborhood characteristics as well as social interactions among probationers. She allows for two types of social interactions: the mean hazard rate for probationer's in i's neighborhood, m_g, and mean time to recidivate in the population, r_g. These are estimated over the entire sample, and so are not time varying. The hazard rate she employs takes the functional form

$$m_{ig}(t, x_i, \gamma_g, m_g, r_g) = \varepsilon_0(t) \exp(k + cx_i + d\gamma_g + J_1 m_g + J_2 r_g), \tag{75}$$

where $\varepsilon_0(t)$ denotes the baseline hazard function. (Since the model is expressed in continuous time, t is treated as an argument rather than a subscript.) Sirakaya addresses unobserved group effects by considering frailty model variations of (75) which helps address issues of unobserved group effects. Probationers are assigned to neighborhoods, which eliminates issues of self-selection. Sirakaya finds strong evidence that endogenous social interactions effects matter.

A particularly sophisticated analysis is due to de Paula (2009) who uses the same data explored by Costa and Kahn (2007), mentioned in the introduction as an outstanding example of the use of history to uncover social interactions, to estimate hazard models of desertions among Union soldiers in the Civil War. de Paula develops a formal model of exit decisions in which the payoff of desertion exhibits dependence on the desertions choices of others. In his model, the expected payoff to desertion at time τ is decreasing in the fraction of the population of prisoners that have previously deserted, in the sense

that the utility of desertion at a given point in time is increasing in the number of contemporaneous deserters, which depends on the size of contemporaneous prisoner population. de Paula formalizes this idea using a continuous time coordination model. Identification is formally proved for the model and evidence is adduced that bunching in desertions occurred in a way that is consistent with social interactions. This work is extended to an elegant discussion of semiparametric identification of interdependences in duration models in Honoré and de Paula (2010).

c. Uncovering social interactions via their effects on laws of large numbers and central limit theorems

When social interactions generate dependence across agent behaviors in a group, their presence will have implications for the convergence rates of sample means and so will affect laws of large numbers and central limit theorems associated with data sampled from the group. A number of authors have proposed ways to exploit these effects in order to generate social interactions. A neglected theoretical predecessor to this social interactions work is Jovanovic (1987) who studies how interdependences could lead idiosyncratic shocks to produce aggregate uncertainty.

One approach of this type is due to Glaeser, Sacerdote and Scheinkman (1996). Their objective is to examine whether endogenous social interactions contribute to cross-city variation in crime rates. One can interpret cities as groups and code the crime/no crime choice as $\omega_i = 0$ and $\omega_i = 1$ respectively, in order to preserve our binary choice notation. If one thinks of persons across all cities as having a common probability p to commit a crime, then the crime rate for the population of city g will have an associated variance of $p(1-p)/n_g$. On the other hand, the presence of social interactions may increase this variance by introducing dependence across choices. To formalize this intuition, Glaeser *et al.* consider a model in which individuals are placed on a line and indexed outwards from the origin, $\{0, \pm 1, \pm 2\}$, so that a city of size $n_g = 2n + 1$ will have individuals ranging from $-n$ to n. They propose a stochastic process for choices in which individuals in a city come in three types: type 0 individuals are always law abiding, type 1's are always criminals, and the remaining type 2's mimic their predecessor in the order. The assignment of types to locations on the line is i.i.d. They show that this model produces greater cross city variance in crime rates than the model without social interactions. Specifically, the variance in the crime rate is $p(1-p)(2-\pi)/\pi n_g$, where p is the probability that an individual with fixed behavior is a criminal, $(1-p)$ is the probability that a fixed individual never commits a crime, and π is the probability that an individual is a fixed type. Without social interactions, $\pi = 1$; that is, everyone is either type 1 or type 2. The presence of a group in the population that can be influenced raises the variance. They propose independently estimating p, and testing for social interactions by comparing the variance of cross-city crime rates with $p(1-p)/n_g$, that would result from no social interaction.

In order to operationalize this approach it is necessary to control for cross-city differences in crime rates that are due to differences in fundamentals. That is, the cross-city test requires that under the null of no social interactions, individuals within and across cities have equal probabilities of committing crimes. Glaeser *et al.* address this both by accounting for observable crime determinants such as demography and by either employing functional forms to allow the fraction of the independents in a population to be estimated in the presence of unobservable fundamental crime determinants or by bounding the importance of the fundamentals "by assuming that the unobservable heterogeneity has twice the predictive power of a one-year lag of the city's own crime rate" (Glaeser, Sacerdote, and Scheinkman, 1996, p. 518).

One limit of the Glaeser *et al.* approach is that evaluation of the role of social interactions requires assumptions that rule out various types of cross-section dependence across agents, in order to construct test statistics with standard distributions. This is a standard problem in cross-section data analysis, since these dependences presumably occur in some sort of social space (Akerlof, 1997) and there is no natural metric for measuring distance between individual observations with respect to which one can argue that distance is attenuated, as can be done in time series and spatial contexts and helps motivate HAC estimators (West, 2008). In fact, Glaeser *et al.*'s reliance on aggregate data rules out the use of such corrections even if the social space were observable.

de Paula and Tang (2010) provide a set for tests of social interactions that may be interpreted as extensions of the Glaeser, Sacerdote and Scheinkman approach to looking at the properties of sample moments. de Paula and Tang consider binary choices in which the payoffs in (63) are modified to

$$V_i(1) - V_i(-1) = k + cx_i + J(x_i)\sum_{j \neq i} \omega_{jg} - \varepsilon_i.$$

There are several qualitative differences with (63). First, the endogenous social interactions parameter is allowed to depend on x_i. The authors are interested in the case where the parameter is negative as well as positive. Second, the endogenous effect depends on the number of agents making the choice, not the average. Third, while the idiosyncratic shocks are still assumed to be conditionally independent, their distribution functions are modeled as $F_{\varepsilon|x}$, so that each distribution function may depend on the individual's characteristics. The information set for agents is assumed to be the same as in Brock and Durlauf (2001a) and elsewhere, and so leads to a Bayes-Nash equilibrium of the type we have studied.

de Paula and Tang argue that even with individual level data, this model is not identified. They therefore propose to study cases where groups are composed of individuals with identical x_i values. This leads them to argue that multiple equilibria in groups with a given $x_i = \bar{x}$ can identify the sign of $J(\bar{x})$. When $J(\bar{x})$ is positive, this is easy to see, since different groups will have different expected average choice levels and so in the Glaeser *et al.* sense

produces excess intergroup variance in sample means. While this was originally recognized in Brock and Durlauf (2001b), de Paula and Tang develop the argument.

Further, de Paula and Tang argue that multiple equilibria can hold when $J(\bar{x})$ is negative. Their identification argument differs from our previous arguments on how multiple equilibria facilitate identification. de Paula and Tang shift their analysis from average choice to individual choices within a group. They show how negative $J(\bar{x})$ can mean that there is a negative correlation among intragroup choices. The key to their analysis is that even though aggregate quantities such as the expected average group choice may be constant, there are multiple equilibria with respect to which agents choose 1 as opposed to -1. This represents a new view of the informational content of multiple equilibria. This approach does require individual level data, unlike Glaeser *et al*. We conjecture that if one focuses on average group behavior, a negative $J(\bar{x})$ would lead to lower variance in the sample averages of group behavior than would occur when social interactions are absent, so that the aggregate approach of Glaeser *et al*., may be applied to test for social interactions for this case as well.

Another approach to the identification of social interactions via qualitative features of sample moments was suggested in the context of financial applications in an early paper by Brock (1993) and uses bifurcations around certain parameter values of a type where the Law of Large Numbers and the Central Limit Theorem break down as in, for example, the statistical mechanics models of Amaro de Matos and Perez (1991) and Ellis (1985). The basic idea of this second approach is to explore how strong dependence between choices can lead to qualitative changes in the properties of the joint stochastic process for a set of choices. These types of breakdowns occur in variations of the binary choice model, and they have some surprising consequences. In the linear in means model, it is natural to use the sample mean as an instrument for individuals expectations. Since equilibrium requires that individuals' beliefs are the correct first moment of the population distribution of choices, this amounts to using the sample mean as an instrument for the population mean. While this approach is well-justified in the linear in means model, it creates an equilibrium selection bias in binary choice models. If $\beta J > 1$, the model has multiple equilibria. Brock and Durlauf (2001b, pp. 3364–7) show that there is a function $H : \mathbf{R} \to \mathbf{R}$ which can be thought of as a potential function. It has the property that if m is a local maximum of $H(\beta Jm)$, then m is an equilibrium expectation. Nonetheless, as the population size grows, $\lim_{n\to\infty} m_g \in \text{argmax } H(\beta Jm)$. Generically, this set is a singleton, and so the procedure of replacing the population mean with the sample mean in effect selects one equilibrium, the equilibrium which globally maximizes $m \mapsto H(\beta Jm)$. The selection of equilibrium by an estimator should appear to be quite troubling. The argument has been made that different dynamical processes of choice revision by individuals (such as best-response and learning dynamics) select the potential-maximizing equilibrium.[51] For the economist

[51] See, for instance, Blume and Durlauf (2003) and Monderer and Shapley (1996).

who is aware of these results, the use of the sample mean as an instrument for beliefs in binary choice models may be a virtue rather than a vice. One implication of this selection effect is that estimates using the sample mean will behave discontinuously in the parameters of the model. The correspondence from model parameters to global maxima of H is upper hemi-continuous but not continuous — small changes in parameters can produce big changes in the location of the global maximum (although not in the maximal value of H).

Brock and Durlauf (2006, section 2.3) extend this type of argument to the multinomial case to show how a tiny change in the distribution $F_{h|g}$ of the characteristics of group g can cause a large change in the limiting value of the fraction of group g choosing choice l among possible choices $0,\ldots,L-1$, provided βJ is greater than some critical value. This approach suggests that for general social interactions structures a potential route to identification would be to estimate the sum of absolute values of correlations among members of a group, denote this S_t and look for dates t^* where S_t changes abruptly. While it is possible that the stochastic structure of the generating processes of unobservables and selection effects that have not been accounted for by estimating a model of selection into groups, e.g., equation (39) could display similar abrupt changes.

We believe that exploitation of potential bifurcations due to endogenous switching from weak dependent data generating processes to strong dependent data generating processes is a potentially important area for future research on the detection of footprints of endogenous social interactions in various data sets. At the very minimum this kind of approach to identification issues should help the scientist to get a better understanding of mechanisms of abrupt change, whatever the ultimate cause of such abrupt change may be.[52]

A different approach to this type of analysis is to explicitly study the implications of dynamic analogs to the discrete choice models we have described. Blume and Durlauf (2003) do this by explicitly considering evolutionary game analogs of the Brock-Durlauf model. They show how the fraction of average time spent around the higher average utility state of the model approaches 1 as noise in the individual choices shrinks to zero, while first mean passage times at the extremal equilibria become unbounded. These sorts of properties represent the dynamic analogs to the use of cross-section properties to infer social interactions. Hence, we also see a need for further study of dynamic social interactions models as these dynamics can produce distinct footprints of social interactions.

d. Beyond Bayes-Nash equilibrium

Very recent work on discrete choice models of social interactions has focused on relaxing equilbrium belief restrictions. One approach is due to Li and Lee (2009) who

[52] See Brock (2006) as well as other articles in Repetto (2006) in the context of trying to explain possible causes of abrupt movements in environmental policy as well as other abrupt movements in policy.

employ an interesting data set on the 1996 Clinton versus Dole Presidential election. In this data set individuals were asked about their own intended vote and whether they thought their reference group members (where the reference group was well defined in the data set) would vote for Clinton, Dole, a third party candidate, or not vote at all; these data are trimmed to produce a binary choice between the two major party candidates. Using these data on the beliefs of each respondent about the voting choices of his reference group, Li and Lee compute a subjective expectation, which they denote by p_{ig}, for each individual i, to play the role of m^e_{ig} in equation (62) above, and use it to test the null hypothesis of equilibrium beliefs by the goodness of fit in the binary choice model with independent types and equilibrium beliefs, as opposed to the subjective expectations they construct. Using maximum likelihood estimation methods, Li and Lee produce two interesting results in this part of their paper. First, they show that estimation effectiveness (measured by the size of the likelihood and in-sample and out-of-sample prediction results) of the binary discrete choice model with social interactions is improved when the subjective expectation data are used in place of the equilibrium beliefs version of the model. Second, they reject the null hypothesis of rational expectations. Incorporation of group level unobservables does not qualitatively affect these findings.

Notice, however, that their findings are still consistent with Bayes-Nash equilibria, because conditional on their own type, people could have different beliefs about the types of others. In Bayes-Nash terms, this rejects the hypothesis that types are independent. From this perspective, one could argue that Li and Lee have rejected the complete (but imperfect) information version of their game in favor of the incomplete information version.

Beyond the testing of the rational expectations hypothesis, Li and Lee make a second methodological contribution by introducing heterogeneity in the social interactions parameter across agents and allowing the data to measure this heterogeneity by estimating a random coefficients version of the model using simulation estimator techniques. They also estimate a threshold version of the model in order to allow for nonprobabilistic voters, i.e., "die hard" Clinton and Dole supporters who vote for their preferred candidate with probability one. They find little evidence of parameter heterogeneity and virtually none for nonprobabilistic voters.[53]

A very general analysis to understand the empirical implications of social interactions when rationality assumptions are relaxed is due to Aradillas-Lopez (2009). This

[53] Li and Lee conduct in-sample and out-of-sample prediction analyses and find that out of sample prediction is improved when endogenous social interactions are included. Thus, regardless of whether the estimate of the social interactions parameter is simply picking up selection effects or group level unobservables, it can still facilitate prediction. However if one wants to make policy statements where it matters whether the endogenous social interactions parameter is estimated consistently or not, then predictive improvement is an insufficient basis for using the model. This is related to Manski's (2010) argument on the usefulness of reduced forms versus structural models of social interactions which is discussed in section 3.

approach is based on Aradillas-Lopez (2008) and Aradillas-Lopez and Tamer (2008) which develop methods for estimating games in which players are sophisticated but not rational in the standard sense.[54] Aradillas-Lopez and Tamer do this by assuming that individual beliefs are rationalizable in terms of iterated-dominance; this allows them to consider identification for different levels of iteration. In contrast, Aradillas-Lopez (2009) focuses on a different notion of rationalizability for his social interactions context, which we now describe.

Aradillas-Lopez (2009) considers a multinomial choice social interactions environment. Individual payoffs are not parametrically specified, as occurred in section 5.vi.a. Rather, the only assumption imposed on these payoffs is that they exhibit supermodularity with respect to the choices of others. The supermodularity-type condition in Aradillas-Lopez (2009) takes the form of an increasing differences assumption of payoffs with respect to an aggregate variable that summarizes the strategic interactions effect. As noted in section 2, supermodularity is a type of endogenous social interaction effect; note that as a property of payoffs, it does not correspond to a parameter analogous to J in our parametric models but rather is defined by a qualitative feature of individual payoffs. Aradillas-Lopez' supermodularity conditions produce a model of ordered choice with social interactions, a class of models that had not been previously studied.

Aradillas-Lopez develops sharp probability bounds for choice probabilities under the assumption of rationalizability. Instead of using an iterated-dominance definition for rationalizability, a set of choices in a population is considered to be *rationalizable* if it represents a set of individual best responses to an associated profile of individual beliefs each of which includes the joint outcome in its support. At the individual level, this requires that each agent assigns a strictly positive probability to the realized choices made by his opponents. These nonparametric probability bounds allow for separate inference properties of individual payoff functions, specifically on whether they depend on the choices of others. When multiple equilibria exist, these bounds can also be used to characterize the selection mechanism used by agents to coordinate on a given set of equilibrium choices. Aradillas-Lopez succeeds in finding empirical restrictions implied by social interactions in an environment with nonparametric payoff functions. His ability to do so under such weak restrictions on beliefs strikes us as remarkable, and additionally represents a demonstration of the power of the supermodularity concept in econometric work.

6. EXPERIMENTAL APPROACHES

This section considers different approaches to the identification of social interactions that involve various forms of experiments, ranging from the laboratory experiments in which the analyst is free to specify much of the socioeconomic environment to

[54] We thank Andres Aradillas-Lopez for much help in writing this discussion.

quasi-experiments in which a change in some environment produces experimental-type conditions to social experiments in which a policy change is implemented in order to generate evidence of social interactions effects.

i. Laboratory experiments

Given the difficulties involved in identification of social interactions from nonexperimental data, laboratory experiments would seem to offer a promising alternative for studying social interactions. It should be possible, for instance, to create experimental designs such that $\bar{x}_g$ does not lie in the span of the elements of γ_g, thus achieving differentiation of contextual from endogenous effects in the linear in means model. Unobserved group characteristics are essentially a measurement problem. By controlling what group members know about each other, and by defining the environment of the interaction, unobserved group characteristics can be eliminated. Finally, group membership can be explicitly controlled, which addresses the self-selection issues.

The experimental literature on social interactions is voluminous. The investigation of social preferences, including such concepts as altruism, fairness, inequity aversion, reciprocity, and trust, is central to experimental economics, and regarded as a significant achievement by its practitioners. Nonetheless, experimental economists have eschewed serious econometric investigation of these data. Rather than bringing the formal tools of statistical inference to experimental data, verbal models and arguments and descriptive statistics are deployed to make arguments from experimental data.

We will argue that despite the increased opportunities for control offered by the experimental laboratory, issues of identification do not disappear. Rather than attempt to survey this literature, we will focus on identification issues that arise in one particular game, the trust game. The trust game was first studied in the laboratory by Berg, Dickhaut, and McCabe (1995). The game is played as follows: A "trustor" is given a stock of money. He may keep all or any part of it for himself, and turns the rest over to the "trustee". The experimenter increases the money handed to the trustee, so he receives a multiple greater than one of the amount entrusted to him by the trustor. He may keep all or any part of the money for himself, and returns the rest to the trustor. At the conclusion of the game, the trustor has received what he kept for himself plus what he received from the trustee. The trustee has received what he kept for himself.

If dollar returns are taken to be equivalent to payoffs, the only Nash (not just subgame perfect) equilibrium of this game is the trustee to return 0, and thus for the trustor to give 0 to the trustee. Deviations from the 0, 0 play are alleged to be indications of trust. Trustors in the Berg *et al.* experiment were given 10 dollars to allocate, and the share going to the trustee was tripled. The average trustee's share was 5 dollars, and only two of the thirty-two subjects they reported on gave 0 to the trustee. The trustees were less generous, returning on average 95% of what the trustor gave. (Only 11 out of

30 trustees returned to the trustor more than he invested.) Thus, the trustors had a negative return on their investment in the trustees. Half of the trustees repaid a dollar or less. These findings are robust. Repetitions of this experiment have been done all over the world, with more or less similar findings. Different versions of the game have been examined as well. Most popular have been binary choice games in which a trustor chooses either to end the game and receive a fixed payoff, or pass the move to the trustee. The trustee in turn chooses either a fixed split of a larger amount, guaranteeing the trustee more than he would have made had he quit, or keeping the entire amount for herself and returning nothing to the trustor.

Berg *et al.* (Berg, Dickhaut, and McCabe (1995), p. 23) ask the question, "Is trust a primitive in models of economic behavior?" The first difficulty in bringing experimental and other empirical methods to bear on trust is that one has to know what one is measuring. Trust has proven to be a highly elusive concept. Many theorists of trust see it as particular. Baier (1986), Hardin (2002) and others see trust as a ternary relation: *A* trusts *B* concerning *C*. Hardin (2002, p. 1) equates trust with "encapsulated self-interest", by which he means a reliance due to reciprocity. Baier (1986, p. 235) argues that trust is reliance on another's good will, rather than reliance on other attitudes and reactions, such as fear or other incentives. It seems that she would reject Hardin's reciprocity definition, since fear of reprisal is distinct from good will. Baier also claims that potential loss is a precondition for trust. "Trust, then, on this first approximation, is accepted vulnerability to another's possible but not expected ill will (or lack of good will) toward one" (Baier, 1986, p. 235). Social capital theorists such as Coleman (1988) and Putnam, Leonardi, and Nanetti (1993), and the many economists who write on the importance of trust for economic growth such as Knack and Keefer (1997), are interested in general trust rather than in the particular.[55] This is not to say that they believe there are no limits to trust, but they theorize about trust in situations where it would be surprising if reciprocity played a significant role. Trust, for them, is more a generalized version of Baier's good will than it is Hardin's encapsulated self-interest.

Identification would seem to be the first fundamental question one asks when confronting a data set: "Is it at all possible to observe a particular phenomenon in data coming from the supposed data-generating process?" But this question presumes an understanding of what constitutes a measure of the phenomenon being examined. The interpretation of trust-game data suffers from a lack of models and even appropriate categories for defining results. Thus Camerer (2003, p. 87) is able to write within the space of a single paragraph and without any apparent irony, "K. . .found Bulgarian students were surprisingly trusting. . .He speculates that Bulgarians are used to trusting among themselves precisely because their trust in authority is so low. E. . .found very

[55] The identification problems we have described for general social interactions models apply *a fortiori* to empirical studies of social capital, see Durlauf (2002) for discussion.

little trust and trustworthiness among Orma herders in Kenya. . .Note that Kenya is considered one of the more corrupt countries in the world, measured by indices of 'transparency,' which guess the extent of bribery, bureaucratic corruption, and black market trade, so it is encouragingly consistent that this simple game shows low levels of trust also." Scholars even disagree on what is being measured. Camerer (2003, p. 85) writes, "Sociologists and psychologists usually object that this game doesn't capture all there is to trust because the two-person one-shot game does not include relationships, social sanctions, communication, and so many other rich features that may support or affect trust. That's precisely the point—the game requires pure trust." Ermisch and Gambetta (2006, p. 12) disagree. "This fear of 'framing' the experiment simply ends up generating indeterminate stimula. Trustworthiness, by contrast, does not exist in the void. There is no such thing as pure trust. It is always trust in someone to do something, e.g., pay their debts, or look after one's children."

The lack of a theory or even a generally accepted definition of trust has not inhibited attempts to measure it. In addition to experimental measurement from behavior in trust games, survey methods have been deployed in the measurement of trust. The World Values Survey asks, for instance, "Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?" Other surveys elicit a scale measurement of the same general kind of question. Unfortunately, survey and experimental measures of trust frequently disagree. Glaeser, Laibson, Scheinkman, and Soutter (2000) and Ermisch, Gambetta, Laurie, Siedler, and Uhrig (2009) find that attitudinal questions about trust do not predict trusting behavior in trust games. Only Fehr, Fischbach, von Rosenbladt, Schupp, and Wagener (2002) reach a different conclusion.

Keeping in mind that the entire enterprise is not well founded, we now turn to some particular identification issues in the analysis of trust game data. It is clear that for the game-form presented in Berg *et al.*, the choices of a significant percentage of trustors and trustees are not Nash play of any game where utility payoffs depend only on money, and do so monotonically. Researchers have suggested a variety of explanations for the choices of trustees, including a fairness norm and reciprocity. Explanations for the choices of trustors include the expectation of trustworthy behavior and altruism towards the trustees. The goal of statistical analysis is to uncover the distribution of such behaviors in the population of trustors and trustees. There are two ways to proceed. One requires that individuals understand the choice environment the experiment puts them in. The analysis of such *experienced players* proceeds by estimating the distribution of types such that the data appear to be the realization of draws from the Bayes-Nash equilibrium distribution of play. The second approach is to assume that players are inexperienced. In this case, the individuals are not obviously optimizers of anything at all. Their behavior is imported from elsewhere. It is commonly believed that the laboratory game induces a particular frame, perhaps by analogy with common experiences. For instance, Henrich *et al.* (2005, p. 798), reporting on the choices of their

inexperienced subjects in ultimatum games, say of their subjects: "Experimental play often reflects patterns of interaction found in everyday life." This approach poses two problems for the experimenter. First, Manski (2002) has observed that without restrictions on beliefs and decision rules other than that they rationalize choice (that is, without any equilibrium restrictions), the distribution of players' types may not be identified. Second, if players' beliefs and preferences come from frames of reference external to the experiment, it does not seem that much will be learned about the world unless the experimenter can either inquire about or control the subjects' frames. Cronk (2007) demonstrates both the importance of frames and how they may be manipulated in trust games played by Maasai tribesmen in Kenya.

Manski (2002) is concerned with the identification of the distribution of preferences and beliefs in a population of game players. He assumes that choice comes from maximization of expected utility, but does not make use of any constraints imposed on beliefs that would normally come from equilibrium conditions. This is entirely appropriate for modeling games played by inexperienced players, but games played by experienced subjects offer more opportunities for identification because the Bayes-Nash and other equilibrium conditions impose cross-player constraints on beliefs. The hope is that when a suitable equilibrium theory is employed, one can do better than Manski suggests. The following two-player, two-strategy version of the trust game demonstrates some of the opportunities and issues that will be encountered.[56]

Player 1, the trustor, has an opportunity to keep 10 dollars or give it to player 2, the trustee. The money is tripled along the way. The trustee then has the opportunity to split the 30 dollars evenly or to keep it all for himself. The trustor's strategies are illustrated in figure 1. The only Nash equilibrium in monetary payoffs is for the trustee to keep everything given to her, and for the trustor to stay Out. We might imagine however, that trustors and trustees both have utility functions exhibiting pro-social preferences. We suppose that the trustees are of two types: *Keepers* care only about monetary payoffs, and so keep all the money. *Sharers* are sufficiently motivated by fairness or reciprocity that they return half of the proceeds. The probability that a given trustee is a sharer is Q. This is a parameter known to the trustors but not to the econometrician, who must estimate it. Trustors also come in types. The utility of going In is the sum of the expected return, $15Q$, and a utility of being altruistic α. That is, $u_o(I) = 15Q + \alpha$. The utility of Out is just the monetary reward, $u_o(O) = 10$. The type distribution for trustors is uniform: $\alpha \sim U[0, A]$. Trustors and trustees are drawn independently from the appropriate type distributions and are matched to play. The equilibrium of this game

[56] Ermisch and Gambetta (2006) study a similar binary action game. Identification arguments are informal. For instance, they wish to distinguish acting fairly from being trustworthy as motives for the neighbor. They choose a nonsymmetric sharing rule because, "Symmetric payments may encourage fulfilling trust for reasons of fairness, rather than because E does what is expected of her. . ." (2006, p. 17). In other words, their identification strategy is to claim that, ipso facto, any sharing that does not result from a 50:50 split is due to trustworthy considerations and not fairness.

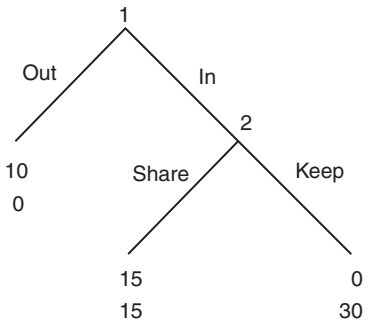

Figure 1 A Trust Game.

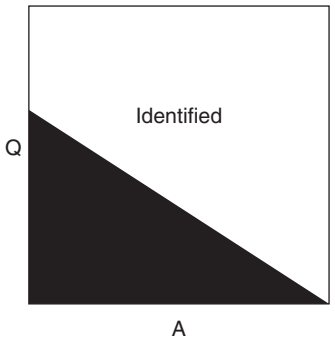

Figure 2 The Parameter Space.

is simple to describe. Sharers Share, and Keepers Keep. Trustors of type $\alpha > 10 - 15Q$ play In, while trustors of type $\alpha < 10 - 15Q$ stay Out. The econometrician wants to estimate the parameters of the type distributions, A and Q. The possible parameters are $(A, Q) \in [0, 10] \times [0, 1]$. Parameter values (A, Q) such that $A + 15Q > 10$ are point-identified. This region is labeled in figure 2. For these parameter values, the fraction of trustors who play In identifies A, and Q is identified by the fraction of trustors who Share. For parameter values on the other side of the boundary, all trustors stay Out, the trustees never get to choose, and their type is never revealed. In this game, it is the action of trustors that allows identification of the trustee type distribution parameter to be observed.

One advantage of the Bayesian framework is that it makes possible inferences across games. For instance, Dufwenberg and Gneezy (2000) consider a variant on the game of figure 1 where trustees can make any division of 20 (rather than 30) should they get the move, and trustors can choose to play In or Out, and receive x, which is varied across treatments. If it is assumed that the type distributions are independent of x, then by changing the treatment, any type distribution can be identified.

One might think that identification problems in this environment stem from the fact that choices have a binary support, which delimits the combinations of behaviors that can be observed. It is instructive to approach the original Berg, Dickhaut, and McCabe (1995) trust game experiment from an econometrics perspective to see how complicated identification can be. Manski's observation, that allowing for rationality but not restricting beliefs through equilibria gives little identifying power, holds here as well. Thus, we model the experiment as a Bayesian game wherein preferences may reflect different motives for trusting and being trustworthy, and see what the data say about the distribution of tastes. We imagine that the subjects of the experiment are drawn independently from populations of trustors and trustees, respectively, and that the Bayesian common prior beliefs are the distribution of players' types in the

population. There are many ways to turn the trust experiment into a Bayesian game. The model we describe was chosen because it suggests some of the statistical techniques that are used in the experimental literature.

To this end, we suppose that trustors are motivated by three potential considerations: The monetary return, conformance to a sharing norm, and altruism. The utility function of a trustor who gives ω_o and receives ω_e from the trustee is

$$u_o(\omega_o, \omega_e) = 10 - \omega_o + \omega_e - \frac{\psi}{2}(\omega_o - \eta)^2 + \rho\omega_o.$$

The first term is the monetary payoff. The second term is the disutility of nonconformity to a social norm η, which is a feasible transfer. The third is altruism, the utility of giving. A type for a trustor is a triple (ψ, η, ρ), and the type space is $T_o = \mathbf{R}_+ \times [0, 10] \times \mathbf{R}_+$. A strategy for the trustor is a function $\sigma_o : T_o \to [0, 10]$.

The trustee has the utility function of the form

$$u_e(\omega_e, \omega_o) = \delta(3\omega_o - \omega_e) - \frac{\phi}{2}(\omega_e - \gamma\omega_o)^2 + \alpha\omega_o\omega_e, \tag{76}$$

where the parameter quadruple $(\delta, \phi, \gamma, \alpha)$ describes the trustee's type. For the utility function described in equation (76), the first term is the monetary payoff, the second is conformity to a social norm, and the third creates a taste for trustworthiness. Notice that the marginal utility of giving depends upon what has been received.

We will simplify our analysis by imposing a constraint on the type space for trustees, that the marginal utility of conformity is positive: $\phi > 0$. Hence utility can be renormalized so that $\phi = 1$. The marginal utility of the transfer ω_e for the trustee when the trustor transfers ω_o is then

$$u'_e(\omega_e, \omega_o) = -\bar{\delta} - \omega_e + (\bar{\gamma} + \bar{\alpha})\omega_o + \nu,$$

where $\bar{\delta}$, $\bar{\gamma}$ and $\bar{\alpha}$ are population means, and

$$\nu = \omega_o(\varepsilon_\alpha + \varepsilon_\gamma) + \varepsilon_\delta.$$

By construction, $E(\nu) = 0$ since the ε's are deviations from population means. The type of a trustee is, given our normalization, the vector (α, γ, δ).

We suppose that individual trustees can differ in their perception of what the norm is, but this is not essential for our analysis.[57] The type space is defined by 1) nonnegativity of

[57] Perceptions of the norm come from the world external to the experiment. If we believed that individuals completely internalized the experiment, then we could impose an additional equilibrium condition on the norm. This belief, however, which we would require for observations of real social phenomena, is unnatural for the lab. Here is an example of how, by not being able to control the frame, the laboratory setting introduces additional noise not present in the world.

the utility parameters, 2) that it is not the norm to give back more than the gross return on the transfer from the trustor, and 3) an a *priori* constraint on the transfer ω_e, that it not exceed the amount of money the trustee has been alloted to divide: $\omega_e \leq 3\omega_o$. Without loss of generality, $\gamma + \alpha - \delta \leq 3$, since behaviors of types with $\gamma + \alpha - \delta > 3$ are indistinguishable from that of types on the boundary, $\omega_e = 3\omega_o$. Thus the type space for the trustees is

$$T_e = \{(\alpha, \delta, \gamma) \in \mathbf{R}^2_+ \times [0, 3] : \alpha + \gamma - \delta \leq 3\}$$

and trustee strategies are functions $\sigma_e : T_e \times [0, 10]$ such that $\sigma_e(t, \omega_o) \leq 3\omega_o$.

The specification of the Bayesian game is completed by specifying type distributions μ_o and μ_e on the type spaces T_o and T_e, respectively, for trustors and trustees. Each individual trustor and trustee knows his own type, and the distribution from which the other type is drawn. In this Bayesian game, the type of the trustor is irrelevant to the trustee since the trustee sees the trustor's action when she must choose. The trustor, however, cannot be certain about how much the trustee will return. The trustor will maximize expected utility, where the expectation will be over the type of the trustee and the trustee strategy function is known to him. The econometrician knows the structure of the game, but sees only transfers ω_o and ω_e. The econometrician's task is to estimate the type distribution, thereby pinning down the relative importance of the different motivations for the transfer of money in the population of experimental subjects.

It is easy to see in this framework how identification problems arise. First, suppose that the type distributions are such that all decisions are interior.[58] The first order condition requires that $u'_e(\omega_e, \omega_o) = 0$, and so

$$\omega_e = -\bar{\delta} + (\bar{\gamma} + \bar{\alpha})\omega_o + \nu,$$

and it is clear that while the marginal rate of substitution between monetary reward and conformity can be identified from the trustee's behavior, the social norm and marginal rate of substitution between altruism or trustworthiness and conformity cannot. One might argue that this is due to the excessive simplicity of the structural assumptions. A more natural assumption might be to assume that the norm is affine rather than linear. This introduces another parameter, and the consequence is that the marginal rate of substitution between private return and conformity can no longer be identified either. Another possibility is to assume, for instance, that the norm is linear while trustworthiness

[58] The set of parameters for which this will be true has a nonempty interior. We do not derive it here, but it is worth noting that sufficient conditions involve both trustor and trustee parameters, since if $\omega_o = 0$, then of necessity $\omega_e = 0$.

is quadratic. This is no less arbitrary than our linear assumption, and leads to a mismatch between norm and equilibrium strategy, which is in some sense more severe than the present model, since the equilibrium strategy would be quadratic in received transfers while the social norm is presumed to be linear.

The trustee's behavior does not exhaust the possibilities for identification. The trustor knows the parameter values that the econometrician does not, and they are payoff-relevant for the trustor's decision. This, unfortunately, does not help. The trustor's optimal strategy is

$$\omega_o = \frac{\rho - 1 + \bar{\gamma} + \bar{\alpha}}{\psi} + \eta$$

(recall that we have assumed that we are in a region of $T_o \times T_e$ where the righthand side is positive), and no additional information is revealed that allows for distinguishing $\bar{\alpha}$ from $\bar{\gamma}$. It is possible however, that variation in the initial stakes provided to the trustor and the rate of return on the transfer to the trustee could lead to additional identifying restrictions on the distributions of both trustor and trustee type distributions.

The constraint that equilibrium is interior is severe, and one might suspect that relaxing it may introduce nonlinearities which may help with identification. Here the blessings are mixed. On the one hand, there is no hope for nonparametric identification, because the conditional means of the parameters given that one is in that set of parameters where the transfer is positive is still not identified. On the other hand, we have checked in several examples that with enough parametric assumptions, point identification can be achieved.

The model just developed addresses particular behavioral hypotheses, and makes particular assumptions about subjects' preferences and behaviors. Other natural behavioral hypotheses can be addressed with different game theoretic models, and these models may offer different opportunities for identification of population types. Levine (1998) for instance, evaluates experimental data for ultimatum games and voluntary contribution games, among others, from the perspective of a model where one player's utility depends explicity on the type of the other rather than just through the other's action. One can see in his games that the hypothesis of equilibrium play and beliefs provides significant identifying power. It should be clear, however, that identification is a nontrivial matter, and that using exercises like this to inform the design of laboratory experiments as well as the analysis of the data, might lead to more convincing experimental research. In particular, experimental design should be informed by what game it has the subjects play, what kinds of preferences they could be expected to hold, what will be determined by the experience they

gain as they learn to play the game, and what is uncontrolled and comes from external frames of reference.

Although the application of equilibrium theory to experimental games holds some promise for the design of economists' identification strategies, laboratory experiments with experienced subjects face additional problems. Chief among them is the problem of external validity. For example, do the results of a private contribution public goods experiment actually say something about charitable giving?

ii. Quasi-experiments

Other authors have focused on changes in group composition whose purpose was not to study social interactions but whose structure is potentially informative of their presence. One well-cited example is Angrist and Lang (2004), which focuses on Boston's Metropolitan Council for Educational Opportunities (METCO).[59] This is a voluntary desegregation program that involves enrolling underprivileged inner city children in suburban public schools. Angrist and Lang (2004) show that the receiving school districts, which have higher mean academic performance than the sending ones, do experience a mean decrease due to the program. However, they also show that the interactions are merely compositional in that there is little evidence of statistically significant interactions of METCO students on their non-METCO classmates. Their analysis with micro-data from one receiving district (Brookline, Massachusetts) generally confirms this finding, but also produces some evidence of negative interactions on minority students in the receiving district. Since METCO is a voluntary program for both sides and thus involves self-selection both at the individual and at the receiving end, at best it can be thought of as uncovering treatment on the treated, which does not translate naturally into claims about social interactions per se for reasons we will discuss in detail below in the next subsection.

iii. Moving to Opportunity

There exists one intervention in group formation that has been implemented on a large scale in order to understand social interactions. Interest in understanding the effects of poor neighborhoods on their residents led the Department of Housing and Urban Development to implement the Moving to Opportunity (MTO) demonstration

[59] Another prominent study of this type is Sacerdote (2001). See Durlauf and Ioannides (2010) for some assessment of its information content with respect to social interactions.

in Baltimore, Boston, Chicago, Los Angeles, and New York, starting in 1994.[60] The program provided housing vouchers to a randomly selected group of families from among residents of high-poverty public housing projects. Within this subsidized group, families in turn were randomly allocated between two subgroups: one which received unrestricted vouchers; and another which received vouchers that could only be used in census tracts with poverty rates below 10% (these users are termed the experimental group). Members of the experimental group also received relocation counseling. The presence of both unrestricted and restricted voucher recipients is a nice feature of the demonstration.

The data from the MTO program have been used to study social interactions in what now amounts to a large number of studies. Goering and Feins (2003) is a compendium of results for the different program sites. Briggs, Popkin, and Goering (2010) is the most recent overview of the demonstration and is of interest not only because of its statistical evidence but because it focuses on qualitative information such as interviews to understand the specific mechanisms associated with program effects. In terms of the scholarly literature, Kling, Liebman, and Katz (2007) represents the most comprehensive analysis both in terms of its methodology and because of its coverage of all the study sites. Earlier analyses include Ludwig, Duncan, and Hirschfield (2001) and Kling, Ludwig, and Katz (2005). As most clearly articulated in Kling, Liebman and Katz, social interactions are inferred from MTO studies when evidence is found that movement of families from high poverty to low poverty neighborhoods is associated with changes in various socioeconomic outcomes. The MTO demonstration has collected data on outcomes ranging from health to education to employment. Focusing on Kling, Ludwig, and Katz, calculations of social interaction effects are derivative from calculations of treatment effects associated with the vouchers. These authors are careful to distinguish between measures of the effects of intent to treat and treatment on the treated, corresponding to the treatment effects associated with eligibility for a voucher and use of the voucher respectively.

[60] One reason why HUD implemented the MTO demonstration was that there was a prior program in the Chicago area that had found large effects from moves from inner city public housing to more affluent suburbs of the city. The Gautreaux program, named after the lead plaintiff in a law suit against the Chicago Housing Authority dating from 1967, led to the movement of some public housing families in Chicago to other parts of the city whereas other families moved to nearby suburbs. Sociologist James Rosenbaum is responsible for the construction of data sets of the families that participated in the Gautreaux program and initiated use of these data to study neighborhood effects. Rosenbaum (1995) is a good overview of Gautreaux findings, which found that families who moved to suburbs exceeded those who stayed in Chicago for a broad range of socioeconomic outcomes. For example, the percentage of college attendees among children who families moved to suburbs was 54% whereas the percentage for children whose families moved to other locations in Chicago was 21%; when one considers only 4-year colleges the attendance rates are 27% and 4% respectively. As Rosenbaum and other students of the Gautreaux data are well aware, there are problems with the data that delimit how informative they are with respect to social interactions. Information about families who moved to suburbs and then returned to the city is missing, so comparisons of city and suburban families at the time the data were collected suffers from self-selection problems. Self-selection was also present in the initial set of families who participated in the program, as the program was restricted to families that had good track records of public housing upkeep. MTO was explicitly designed to avoid these problems.

Overall, Kling, Liebman, and Katz (2007) do not find that residence in low poverty neighborhoods had particularly strong effects on families. This is so even though the housing voucher randomization resulted in families with similar observable characteristics living in very different neighborhoods. For adults, there is little evidence that movement to low poverty neighborhoods affected economic self-sufficiency as measured across several criteria. The strongest effects suggest that lower poverty neighborhoods are associated with better mental health. For children, lower poverty neighborhoods appear to benefit teenage girls for outcomes ranging from education to risky behavior to physical health. In contrast, teenage boys seem to be adversely affected by relocation to a lower poverty neighborhood. While there was deterioration along a number of outcomes, increases in criminal behavior by the teenage boys was particularly noteworthy. While other authors are more sanguine in their reading of the findings from the MTO demonstration, it is difficult to conclude from the body of empirical work that the movement of poor families to more affluent neighborhoods did much to improve overall socioeconomic outcomes.

Durlauf and Ioannides (2010) discuss some of the difficulties in drawing policy conclusions from the MTO data. For the purposes of this chapter, the relevant question is the extent to which the MTO findings speak to the empirical salience of social interactions. While at first glance, it would seem the unimpressive evidence of treatment effects from the low poverty neighborhood vouchers speaks against the importance of neighborhood social interaction effects, such a conclusion is questionable. The basic problem concerns the mapping from treatment effects that have been estimated from the MTO data and social interactions mechanisms. Moving a family from a high to low poverty neighborhood means that a range of neighborhood characteristics beyond the aggregate poverty rate are changed. For example, while low poverty neighborhoods might provide "better" role models from high poverty ones, they will also be associated with weaker social ties, friendships, proximity to relatives and the like. Movements from high to low poverty neighborhoods cannot be facilely equated with movements from worse to better neighborhoods. These moves represent changes in a vector of neighborhood characteristics that constitute types of social interactions. There is no reason to think that the moves observed in the MTO study did not constitute a tradeoff of some elements of this vector for others.

This problem does not merely provide a reason why certain treatment effects can be small in the presence of large social interaction effects, but indicates why the interpretation of strong treatment effects as evidence of social interactions is problematic. One example is discussed by Kling, Liebman and Katz: the reduction in asthma among children who move to low poverty neighborhoods. One reason this may occur is due to stress reduction. To the extent this is generated by neighborhood characteristics this is reasonably a social interaction effect. Another reason for lower asthma rates, however, could be reduced exposure to vermin infestations, which in turn could be a

consequence of changes in housing quality induced by the requirement that the vouchers were used in low poverty neighborhoods. This is not a social interactions effect.

These considerations suggest that an important next step in the study of MTO data is an explicit consideration of how the data can be directly employed to evaluate social interaction effects. This may involve the development of measures of "social interactions-relevant treatment effects" which contrast with the standard treatment effect measures that economists use; such a development would parallel the work on policy-relevant treatment effects of Heckman and Vytlacil (2001). For example, the issue of self-selection in voucher use (less than a quarter of eligible families took up the restricted voucher) is a nuisance from the perspective of some treatment effect calculations, but may be informative in terms of social interactions, for reasons we have discussed in the context of the linear in means model and elsewhere.[61] These suggestions are admittedly speculative, though we note that this is the flip side of the analysis in Manski (2010) who argues that certain types of treatment effects may be identified without full knowledge of social interactions structure.

To be clear, the body of existing MTO studies is of great importance in thinking about alternative policies. One example that is suggested by the findings in Kling, Liebman and Katz is the distinction between efforts to promote socioeconomic integration of neighborhoods versus effects to reduce poverty rates in the least affluent neighborhoods. Kling, Liebman, and Katz find that changes in neighborhood poverty rates are responsible for those MTO voucher effects they do find as opposed to moves per se, hence the suggestion that it might make more sense to lower poverty rates in currently poor neighborhoods than to promote general socioeconomic integration. Our criticisms only apply to the extent to which claims about the significance of social interactions are made based upon the calculations that have appeared. To return to our argument, the finding that, conditional on poverty rate changes, moves had no positive effects does not speak to the magnitude of social interaction effects that are uncorrelated with poverty levels, which presumably included social features such as friendship ties that were damaged by the moves to more affluent communities.

7. SUGGESTIONS FOR FUTURE DIRECTIONS

In this section, we suggest some new directions we regard as promising in developing a full econometrics of social interactions.

i. Measurement

The empirical literature on social interactions suffers from serious measurement problems. This is a first important area that needs new econometric work. Here we follow the discussion in Durlauf and Ioannides (2010). Economic theory does not

[61] This may suggest new uses for the Gautreaux data as well.

dictate the appropriate empirical measures of contextual variables that a researcher ought to use. As a result, one for example finds Bertrand, Luttmer, and Mullainathan (2000) using the product of welfare usage and own-ethnic group intensity to explain individual welfare usage, whereas Aizer and Currie (2004) use the utilization rate of an individual's language group to measure social interactions on public prenatal-care utilization. Similarly, the empirical literature does not typically consider how social variables should interact with individual decisions, so that linearity assumptions are too often employed without reflection. If the reason why utilization of social services depends on the usage of others is because of information transmission, as argued by Bertrand *et al.*, then it is unclear why the percentage of users is the appropriate variable, as opposed to some nonlinear transformation, as presumably one only needs one neighbor to provide the information. While considering this type of problem in studying social interactions in marriage markets, Drewianka (2003) argues that a higher marriage rate in a community may reduce the propensity of unmarried people to marry as a higher rate hampers search.

A second measurement problem arises because theory does not provide guidance as to the appropriate measure of groups. One aspect of this problem is relatively simple and involves the choice of scale for a given measure of group; in the case of physical proximity, one finds the use of Zip codes in Corcoran, Gordon, Laren, and Solon (1992) and census tracts and blocks in Weinberg, Reagan, and Yankow (2004) to determine neighborhoods. In other cases, measurement problems involve the categories that define groups. The study by Conley and Topa (2002) is unusual in seeking to identify the appropriate axes on which to situate the actors in social space, finding that ethnicity is of particular importance in defining social interactions for labor market outcomes in Chicago. To do this, Conley and Topa construct measures of proximity between neighborhoods based upon physical proximity, ethnic similarity and socioeconomic similarity to determine which measure best explains correlations in unemployment patterns across neighborhoods, finding that ethnicity dominates the other measures.

Conley and Topa's work is an important advance, but follows the literature in equating group memberships with scalar categories. Typically, categories are treated as the relevant groups for social interactions. Hence one sees one body of research that explores the effects of the residential neighborhood in which a person grows up on his adult economic prospects while another explores the effects of friends' smoking behavior on individual smoking decisions. Akerlof (1997) argues that social interactions are best understood as occurring in a social space that may have many dimensions; this follows naturally when one considers the overlapping interactions of factors such as physical proximity, ethnicity, gender, and education on the ways in which individuals interact. One limitation of much of the existing literature is thus the absence of attention to the multiple groups to which an individual is a member.

The major exception to this claim is the networks literature, which allows for differential social connections within a population, but even here, the analysis does not focus on how categories interact to produce an individual-specific social environment.

The development of methods for dealing with the proper measurement of social space represents one of the key outstanding research areas in the econometrics of social interactions. A few papers have explicitly attempted to explore multiple channels for social interactions. A nice example is Iyer and Weeks (2009) who consider the joint roles of ethnicity and geographic proximity in fertility decisions in Kenya, finding a particularly strong role for ethnicity. Methodologically, Conley and Topa (2003) is unique in considering the identification of social interactions when groups are measured with error and when the relevant group for social interactions is not observable, but some other grouping which contains the behaviorally relevant group is observable, as may occur if an analyst knows an agent's census tract but not his actual neighborhood. They find that measurement error leads to nonidentification. Simulation evidence suggests that local identification may still be possible when superset data are available. As Conley and Topa acknowledge, their analysis applies to a very specific model, so much more research is needed on this question.

Some aspects of measurement problems relating to groups have been addressed through the development and clever analysis of interesting data sets. An especially nice example is Mas and Moretti (2009), who employ a data set that measures supermarket employee productivity in 10-min intervals. Their data set is also impressive because the set of peers for a given worker regularly changes due to differences in shift composition and because the spatial orientation of workers in a store is known. This allows for analyses of such questions as whether frequent interactions induce stronger social interactions and whether physical proximity to others matters. Guryan, Kroft, and Notowidigdo (2009) follow a similar strategy using data on performance in golf tournaments but find little evidence of social interactions. These types of data sets, in our view, provide clean ways of describing social space, but are of course limited in terms of the extent to which their findings may be extrapolated to alternative environments.

ii. Social interactions and prices

Most social interactions work has ignored the informational content of prices for group membership. For example, social interactions of residential neighborhoods will be reflected in housing prices, via standard hedonic price arguments. Nesheim (2002) is a pioneering advance in this regard. Social interaction effects, measured as averages of parental characteristics, can be extracted from housing prices using hedonic pricing

methods. Implementation of Nesheim's approach is facilitated when assumptions are made to allow an explicit solution for the hedonic price in terms of neighborhood characteristics; see Ioannides (2008) for a very straightforward way of doing this. So far, Nesheim's methods have not received the empirical attention they warrant. Bayer, Ferreira, and McMillan (2007) report some nonstructural hedonic regressions of housing prices on neighborhood characteristics. See also Bayer and Ross (2009) who propose using neighborhood prices to construct a control function to proxy for unobserved neighborhood characteristics.

iii. Group characteristics as evidence of social interactions

Another dimension along which endogenous group formation can be used to provide evidence of social interactions is the equilibrium distribution of types across groups. The informational content of this distribution was first recognized in the context of racial discrimination in Becker (1971). Becker showed that taste-based discrimination may not manifest itself in black white wage differences but rather in segregation of a subset of firms. Analogous reasoning applies in social interactions contexts. Models such as Bénabou (1993, 1996) and Durlauf (1996a,b) emphasize how social interactions can produce stratification of neighborhoods by income; work such as Epple and Sieg (1999) and Calabrese, Epple, Romer, and Sieg (2006) show how these types of effects can be incorporated into sophisticated models of locational choice; the latter paper is of particular interest since social interactions are essential to the analysis. Yet another context where group compositions are informative about interactions concerns assortative matching, where as discussed above, following Becker (1973), supermodularity in production functions can produce efficient stratification of firms by ability. In general, the tight relationship between supermodularity and stratification has been underutilized as a strategy for uncovering social interactions.[62] We believe that group composition represents a potentially powerful source of evidence on social interactions.

Card, Mas, and Rothstein (2008) can be interpreted as using group composition to uncover social interactions. The objective of their analysis is to assess the empirical salience of the classic Schelling (1971) tipping point model of segregation. Schelling's analysis assumes that there is a threshold value for the percentage of blacks in a community such that whites will move when the threshold is exceeded, generating rapid segregation. Card et al. study the time series of neighborhood racial compositions across the United States and find evidence of break points in the series that are consistent with the predictions of the Schelling model. In our view, this represents partial identification evidence on the effects of neighborhood racial composition on individual utility. What is

[62] See Durlauf and Seshadri (2003) and Prat (2002) for examples of a tight supermodularity/stratification link for payoff functions other than those studied by Becker as well as for some caveats.

important from the perspective of our discussion is that the Card *et al.* evidence derives entirely from group composition.[63]

iv. Joint modeling of group memberships and behaviors

Our discussion of prices and group characteristics as sources of information on social interactions suggests yet another direction for new research: the joint modeling of group memberships and behavioral choices as facets of a general decision problem. Brock and Durlauf (2006) give an example of this perspective using a logit framework; we borrow heavily from their original presentation. The basic idea of this approach is to model individuals as making joint choices of group memberships, $g \in \{0,\ldots,G-1\}$, and behaviors, $l \in \{0,\ldots,L-1\}$. Group choices are denoted as δ_i while ω_i continues to denote the behavioral choice. This joint decision is sequential as groups are chosen first and then behaviors are chosen once groups form; this particular sequencing renders the model mathematically equivalent to a standard nested logit model (Ben Akiva (1973) and McFadden (1978)) with the exception of the presence of endogenous social interactions.

The sequential logit structure ensures that choice probabilities at both stages have a multinomial logit probability structure. Defining $h_{ilg} = k_l + c_l x_i + d_l \gamma_g$, the behavioral choices conditional on a group choice g will be defined by the probabilities

$$\mu\left(\omega_{ig} = \bar{i} \,|\, (h_{ilg}, p^e_{ilg})_{l=0}^{L-1}\right) = \frac{\exp \beta(h_{ilg} + Jp^e_{ilg})}{\sum_m \exp \beta(h_{img} + Jp^e_{img})}. \tag{77}$$

Group choices reflect the fact that choices in the stage will produce utility in the fashion of our original multinomial choice model. This is operationalized by making the group choice probabilities depend on the expected utility of the choice ω_i will produce in the second stage. Letting $\delta_i = g$ code for the choice of group by individual i, these choices are also assumed to exhibit a logit structure:

$$\mu\left(\delta_i = \bar{g} \,|\, (h_{ilg}, p^e_{ilg})_{l=0}^{L-1 \;\; G-1}_{\;\; g=0}\right) = \frac{\exp \beta_G Z_{i\bar{g}}}{\sum_g \exp \beta_G Z_{ig}} \tag{78}$$

where β_G denotes the heterogeneity parameter for the unobservable shocks associated with group choices and

$$Z_{ig} = E\left\{ \max_l \left(h_{ilg} + Hp^e_{ilg} + \varepsilon_{ilg}\right) \,|\, (h_{ilg}, p^e_{ilg})_{l=0}^{L-1} \right\}.$$

[63] The nonlinear model estimated in Easterly (2009), which leads him to argue that the Schelling model does not explain the global dynamics of segregation in the US, is also consistent with the presence of social interaction effects in terms of preferences over neighborhood racial composition, although it is not clear that the model provides affirmative evidence of these effects that is as strong as found in Card *et al.*

Implicit in equation (78) is the existence of unobservable location-specific utility terms that are irrelevant with respect to the utility of a choice once the group is formed. A standard result is that[64]

$$E\left\{ \max_{l} \left(h_{ilg} + Hp^{e}_{ilg} + \varepsilon_{ilg}\right) | \left(h_{ilg}, p^{e}_{ilg}\right)_{l=0}^{L-1} \right\} = \beta^{-1} \log\left(\sum_{i} \exp \beta\left(h_{ilg} + Jp^{e}_{ilg}\right)\right). \quad (79)$$

Equation (79), together with equations (77) and (78) produce a joint probability description of group memberships and behaviors

$$\mu\left(\omega_{i\bar{g}} = \bar{l}, \delta_i = \bar{g} \,|\, \left(h_{ilg}, p^{e}_{ilg}\right)_{l=0\ g=0}^{L-1\ G-1}\right) = \frac{\exp\left(\beta_G \beta^{-1} \log\left(\sum_l \exp \beta\left(h_{il\bar{g}} + Jp^{e}_{il\bar{g}}\right)\right)\right)}{\sum_g \exp\left(\beta_G \beta^{-1} \log\left(\sum_l \exp \beta\left(h_{ilg} + Jp^{e}_{ilg}\right)\right)\right)} \cdot \frac{\exp \beta\left(h_{i\bar{l}\bar{g}} + Jp^{e}_{i\bar{l}\bar{g}}\right)}{\sum_l \exp \beta\left(h_{il\bar{g}} + Jp^{e}_{il\bar{g}}\right)}. \quad (80)$$

As is well known, the compatibility of the nested logit structure (80) with an explicit utility maximization problem requires conditions on the parameters.[65] One condition that ensures compatibility with a well posed maximization problem is $\beta_G \leq \beta$ (McFadden, 1978, pp.86–7). This condition requires that the dispersion of random payoff terms across groups is greater than the dispersion in random payoff terms across behavioral choices within a group. We close the model with the equilibrium condition: For all i, l and g, $p^{e}_{ilg} = p_{lg}$, which links the choices at the two levels.

Models of this type have yet to be analyzed either in terms of their theoretical properties or in terms of identification. We believe models of this type can prove to be a valuable complement to existing social interactions models. For example, the sequential choice structure introduces a new mechanism by which multiple equilibria may emerge, namely the influence of beliefs about group behaviors on group memberships, which reciprocally will affect behaviors. The presence of these types of multiple equilibria can, we believe, provide information on identification. More generally, the sort of probability structure we have described integrates group memberships and individual behaviors in a way analogous to what we referred to as the Becker view of discrimination, in which wages and firm composition are affected by the existence of taste-based discriminators. Our strong conjecture is that this joint modeling will produce new routes to identification.

This discussion also suggests that work to integrate the evolution of social networks with choices in networks may produce valuable insights. Copic, Jackson, and Kirman (2009) is a rare example in economics of the development of likelihood methods to

[64] See, for example, Anderson, de Palma, and Thisse (1992, p. 46).
[65] See McFadden (1978) and Börsch-Supan (1990).

describe network structure. The sort of coupling we have described for group and behavior choices in the nested logit context should be explored using this type of approach.

v. Transition dynamics versus steady-state behavior

A final research direction we believe can prove to be important concerns the use of transitional behavior to uncover social interactions. The linear in means and discrete choice models look at steady state behaviors in the sense that these systems, including their dynamic analogs, conceptualize the data as drawn from their associated invariant measures. While the duration models we describe, especially that of de Paula (2009), focus on transitions in a population, there has yet to be much systematic exploration of the evidence on social interactions that may be found in transitional dynamics versus steady state behavior. By analogy, the steady state distribution of disease rates across locations in a region will not speak to the contagion mechanism for the disease in the way that would data on the transition of the disease across locations.

Brock and Durlauf (2010) provide an example of how transition dynamics can produce evidence of social interactions. They consider a population of perfect foresight actors who, in continuous time, are deciding whether to adopt a new technology. The cost of the new technology is falling over time. The payoff to adoption is the present discounted value of payoff from the time of adoption. The payoff to adoption is increasing in the fraction of the population that has adopted. Agents are indexed by a scalar x, in which the payoff function is strictly increasing. Data are restricted to $q(t)$, the adoption curve for the technology, i.e., the fraction of the population that has adopted as of time t, and f_x, the cross sectional density of x. Brock and Durlauf consider the case where the distribution of types among adopters at each t is unknown as well as the case where the distribution of types among adopters is known at each t.

From the perspective of steady state behavior, there is nothing that can be learned about social interactions; the steady state data will consist of an adoption rate for the population as well as a cutoff value x such that agents with $x_i < x$ have not adopted while other agents have adopted. Such an observation is fully consistent with individual payoffs being independent of the adoption decisions of others. However, the full adoption curve, which represents the transition dynamics for the steady state adoption rate, can be informative about social interactions. For example, even if f_x is unobservable, $q(t)$ can be informative about social interactions. Brock and Durlauf show that if one is willing to assume that f_x contains some mass points, discontinuities in $q(t)$ can only occur because of endogenous social interactions. This does not follow because of multiple equilibria as occurred in the partial identification results under primary choice. Rather it follows from the fact that since higher x types who have not yet adopted always have a greater incentive to adopt than lower x types, even though the population fraction of the lower x types is larger, self-consistent bunching can

occur at particular t values; intuitively, at these jump points the lowest x will meet the first order condition for adoption with equality whereas all others do not.[66]

Brock and Durlauf (2010) also consider a case where x is a vector consisting of an observed variable x_1 and an unobserved variable x_2. For conditional adoption curves $q(t|x_1)$, social interactions can produce pattern reversals where lower x_1 types adopt before higher x_1 types. A condition such as stochastic dominance of the conditional density of x_2 given x_1 is needed for this type of observation to represent evidence of social interactions. As such, the Brock and Durlauf results are another example of how delineation of a complete economic environment can allow for partial identification of social interactions under what appear to be modest assumptions. Arguments of their type can be taken further, as is done in Young (2010), which we discuss next.

vi. Microfoundations

A final area that warrants far more research is the microfoundations of social interactions. In the econometrics literature, contextual and endogenous social interactions are defined in terms of types of variables rather than via particular mechanisms. This can delimit the utility of the models we have, for example, if the particular mechanisms have different policy implications. Put differently, the current generation of social interactions models focuses on a relatively crude division of social interactions between factors that are predetermined and those that are contemporaneous; while one can rationalize this division as structural, this is only true by assumption; work in evolutionary game theory, for example, has a much more subtle view of how endogenous interactions arise. Young (2010) is an important next step in the social interactions research program as it explicitly studies the different empirical implications of alternate social interactions mechanisms. Young derives implications for aggregate behavior by considering where the social interaction comes from. Behavioral economists may be interested in individual behaviors for their own sake, but Young demonstrates here that particular features of the process generating the social interaction determine aggregate behavior, and it raises the prospect that microeconomic and behavioral hypotheses about where social interactions come from may be identifiable from aggregate data.

Young writes on identification of types of social interactions in diffusion processes from a theoretical perspective. He examines different diffusion models in a rather general large-population setting. The different explanations, inertia, contagion, social influence and social learning, are sets of assumptions about individual behaviors. The outcome of the analysis is a set of distinct properties of the diffusion curve, a system aggregate, an

[66] The possibility that social interactions could induce discontinuity in adoption curves was first recognized in a relatively neglected paper by Cabral (1990). Differences in the economic environments studied by Cabral and Brock and Durlauf are discussed in the latter. That said Cabral has priority in discovering the qualitative finding.

emergent property of the system.[67] Inertia is the hypothesis that individuals learn privately, but delay in making their decision. He supposes that each individual i in the (continuum) population can be characterized by a switch rate, λ_i, which is independent of the numbers and identities of those who have already switched. Young shows that no matter the distribution of the λ_i, the adoption curve must be concave. Contagion is a process wherein a given individual adopts when she sees an instance of the innovation, or hears about it. Perhaps the most famous model of this kind is the Bass (1969) model of new product adoption. In these models the instantaneous rate at which an individual adopts will depend upon the size of the pool of current adopters. Adoption curves derived from contagion models will be *S*-shaped, and under some reasonable assumptions, it must decelerate when the pool of adopters exceeds 1/2 the population. Social influence models are threshold models. Each individual has a threshold r_i. If the adopting pool contains fraction r_i of the population, then individual i will adopt. Under some mild assumptions, the adoption process either initially decelerates or it accelerates at a super-exponential rate over some time interval.

There are many ways to build social learning models, and Young's main point is that details matter. Nonetheless, for an interesting class of social learning models he shows that diffusion must initially decelerate, but that the rate of diffusion relative to the pool of potential adoptees increases. Under some additional assumptions, Young can show that if the process begins to accelerate, it does so at a super-exponential rate. Thus Young's work suggests strategies for nonparametric identification of the different behavioral models he considers. What happens when the different behavioral models are mixed in the population, and how to build empirical models that would reflect Young's identifying structural features, are still open questions.

8. CONCLUSIONS

As this chapter has demonstrated, a wide range of identification strategies for uncovering empirical evidence of social interactions are available to empirical workers. These approaches range across linear and nonlinear models, cross-section and time series data, and involve a remarkably broad range of portfolios of assumptions. The existing set of identification results thus does not lend itself to any straightforward summary. Rather, the body of arguments we have described represent different approaches to producing evidence of social interactions at two levels. First, under "ideal" assumptions with respect to unobserved heterogeneity, identification questions revolve around the disentangling of types of social interaction effects: contextual versus endogenous. Second, under more realistic specifications of unobserved heterogeneity, i.e., grouped individual-level

[67] Emergent properties of dynamical systems are properties or structural features that occur on scales of aggregation or temporal scales which are different from those of the rules defining the system. See Blume and Durlauf (2001) for discussion of emergence in socioeconomic environments.

heterogeneity as emerges from endogenous group formation and group-level heterogeneity that is not related to social interactions, identification involves the question of whether any evidence may be adduced for social interactions, let alone whether the specific type of social interaction is recoverable from the observed data.

One way to understand the many methods we have described is that they represent points along an "assumptions/possibilities" frontier. As is true throughout economics, there is a tradeoff between the strength of assumptions made prior to empirical analysis and the precision of the empirical claims that follow. And the types of assumptions we have described, whether they represent restrictions on the probability structure of unobservable stochastic processes or substantive assumptions about individual behavior, can never be expected to hold literally. This should not jaundice the consumers of empirical work on social interactions any more than it should affect consumers of other types of empirical social science. Scientific progress, arises from the interaction of *a priori* beliefs, data and logical reasoning. We therefore regard the interplay of economic theory, econometrics and empirical work as all necessary ingredients in understanding the social determinants of individual behavior.

> *It is no disparagement, therefore, to the science of Human Nature, that those of its general propositions which descend sufficiently into detail to serve as a foundation for predicting phenomena in the concrete, are for the most part only approximately true. But in order to give a genuinely scientific character to the study, it is indispensable that these approximate generalizations, which in themselves would only amount to the lowest kind of empirical laws, should be connected deductively with the laws of nature from which they result; should be resolved into the properties of the causes on which the phenomena depend. In other words, the science of Human Nature may be said to exist, in proportion as the approximate truths, which compose a practical knowledge of mankind, can be exhibited as corollaries from the universal laws of human nature on which they rest; whereby the proper limits of those approximate truths would be shown, and we should be enabled to deduce others for any new state of circumstances, in anticipation of specific experiences.*
>
> ***John Stuart Mill,* A System of Logic *(1859)***[68]

[68] *Collected Works of John Stuart Mill*, J. Robson ed. Indianapolis: Liberty Fund Press. Book VI, chapter iii, section 2, pp. 847-848.

A1. DERIVATION AND ANALYSIS OF EQUILIBRIA IN THE LINEAR IN MEANS MODEL

i. Structure

The linear-in-means model can be derived simply as the unique Bayes-Nash equilibrium of a game in which each individual's choice is determined by a private benefit and a conformity benefit. Not surprisingly, the utility functions are quadratic, and the conformity benefit is modeled as linearly decreasing in the quadratic deviation of an individual's choice from the average behavior of all other players. Individuals belong to a common group g of size n_g. Group membership is exogenous. An individual's realized utility depends upon his own choice and the choices of others. Preferences are expected utility, and are of the form

$$u_i(\omega_{ig}, \omega_{-ig}) = \theta_{ig}\omega_{ig} - \frac{\omega_{ig}^2}{2} - \frac{\phi}{2}\mathrm{E}((\omega_{ig} - \bar{\omega}_{-ig})^2), \tag{81}$$

where $\bar{\omega}_{-ig} = (n_g - 1)^{-1} \sum_{j \neq i} \omega_{jg}$ is the average choice of the others in g. The individual marginal benefit θ_{ig} can be linearly decomposed as follows:

$$\theta_{ig} = \chi_0 + \chi_1 x_i + \chi_2 \gamma_g + \varepsilon_i + f_g \tag{82}$$

where x_i and ε_i are observable and unobservable individual characteristics, and γ_g is a vector of observable group characteristics and f_g is a group characteristic observable to all individuals in the group but unobservable to the econometrician. The determination of individual choices is a game of incomplete information, since each individual, and only that individual, observes ε_i. (Group characteristics unobservable to individual group members are irrelevant to choices as this model exhibits certainty equivalence in individual choices.) The ε_i elements are i.i.d. draws from a distribution on the real line $\mathbf{R}$ with mean 0. For expositional purposes it will be useful to write $\theta_{ig} = \gamma_i + \gamma_g + \varepsilon_i$ where $\gamma_g = \chi_0 + \chi_2 \gamma_g + f_g$ is the internally (to the actors) observable group contribution to the marginal utility of ω_i, and $\gamma_i = \chi_1 x_i$, the externally (to the econometrician) observable contribution to marginal utility of an individual's characteristics.

ii. Existence of equilibrium

In a Bayes-Nash equilibrium, each individual maximizes expected utility, taking the expectation on $\bar{\omega}_{-ig}$ with respect to his belief distribution, and all belief distributions will be correct. The first-order condition for individual i is

$$\gamma_g + \gamma_i + \varepsilon_i - \phi(\omega_{ig} - E\bar{\omega}_{-ig}) - \omega_{ig} = 0,$$

and so

$$\begin{aligned}\omega_{ig} &= \frac{1}{\phi+1}\gamma_g + \frac{1}{\phi+1}\gamma_i + \frac{\phi}{\phi+1}E\bar{\omega}_{-ig} + \frac{1}{\phi+1}\varepsilon_i \\ &= \frac{\chi_0}{\phi+1} + \frac{\chi_1}{\phi+1}x_i + \frac{1}{\phi+1}f_g + \frac{\chi_2}{\phi+1}\gamma_g + \\ &\quad \frac{\phi}{\phi+1}\mathrm{E}(\bar{\omega}_{-ig}) + \frac{1}{\phi+1}\varepsilon_i\end{aligned} \tag{83}$$

This equation justifies (6) when there is no group level unobservable and (30) when there is such an unobservable, since the coefficients in (6) and (30) are proportional to those in (83), assuming that an equilibrium exists. Notice that the shock in (83), $(\phi+1)^{-1}\varepsilon_i$, has a variance that is affected by the strength of the conformity parameter.

We find an equilibrium by positing a functional form with undetermined coefficients, and then solving for the coefficients to make the beliefs correct. It will be convenient to define $\bar{\gamma}_{-ig} = (n_g - 1)^{-1}\sum_{j\neq i}\gamma_j$ to be the mean observable type component in the population. This is simply a sample mean. We suppose that for each individual j,

$$\omega_{jg} = A\gamma_g + B\gamma_j + C\bar{\gamma}_{-jg} + D\varepsilon_j + F. \tag{84}$$

We derive consistency of beliefs by assuming all individuals other than individual i are choosing according to this functional form, computing the best response for individual i, seeing that it is of this linear form, and then solving for the coefficient values such that A through F are common through the entire population. We compute the best response simply by deriving an expression for $\bar{\omega}_{-ig}$ by substituting from equation (84) into equation (83). After some algebra one can show that the coefficients in (84) must fulfill

$$A = \frac{1+\phi A}{\phi+1},\ B = \frac{1+\dfrac{C}{n_g-1}}{\phi+1},\ C = \frac{\phi\left(B + \dfrac{n_g-2}{n_g-1}C\right)}{\phi+1},$$
$$D = \frac{1}{\phi+1},\ F = \frac{\phi F}{\phi+1}.$$

Solving these equations gives the values of the undetermined coefficients. Thus

$$\omega_{ig} = \gamma_g + \frac{n_g - 1 + \phi}{(\phi+1)n_g - 1}\gamma_i + \frac{\phi(n_g-1)}{(\phi+1)n_g - 1}\bar{\gamma}_{-ig} + \frac{1}{\phi-1}\varepsilon_i. \tag{85}$$

When the population size is large, this is approximately

$$\omega_{ig} = \gamma_g + \frac{1}{\phi+1}\gamma_i + \frac{\phi}{\phi+1}\bar{\gamma}_g + \frac{1}{\phi-1}\varepsilon_i \tag{86}$$

where $\bar{\gamma}_g$ is the group-level average of γ_i.[69] Recalling the definitions of the γ terms,

$$\omega_{ig} = \chi_0 + \chi_2\gamma_g + f_g + \frac{\chi_1}{\phi+1}x_i + \frac{\phi\chi_1}{\phi+1}\bar{x}_g + \frac{1}{\phi+1}\varepsilon_i$$

where $\bar{x}_g$ is the group mean of the individual characteristics. This expression corresponds to the reduced form equation (11) in the text when there is no group-level unobservable. Extending the model in this and a variety of other ways to match the other specifications discussed in section 3 is straightforward.

iii. Uniqueness of Bayes-Nash equilibrium

A strategy for player i is a map $f_i(\gamma_g, \gamma_i, \gamma_{-ig}, \varepsilon_i) \mapsto \mathbf{R}$, where $\gamma_{-ig} = (\gamma_j)_{j\neq i}$. The preceding section demonstrates the existence of a symmetric Bayes-Nash equilibrium with linear strategies. Discrete-choice models of social interaction are replete with multiple equilibria, so one might believe that multiple equilibria may arise here as well. This is not the case.

Theorem A.1. Uniqueness of equilibrium in the linear in means model. *The Bayes-Nash equilibrium strategy for the model* (81) *and* (82)*, and defined by* (83)*, is unique.*

Proof. Equation (83) implies that $(f_1^*, \ldots, f_{n_g}^*)$ is a symmetric Bayes-Nash equilibrium if and only if for all i,

$$f_i^*(\gamma_g, \gamma_i, \gamma_{-ig}, \varepsilon_i) = \frac{1}{\phi+1}(\gamma_g + \gamma_i + \varepsilon_i) + \frac{\phi}{\phi+1}\frac{1}{n_g-1}\sum_{j\neq i}\mathrm{E}\left(f_j^*(\gamma_g, \gamma_j, \gamma_{-jg}, \varepsilon_j)\right)$$

Let B_i denote the set of measurable functions f_i: $(\gamma_g, \gamma_i, \gamma_{-ig}, \varepsilon_i) \mapsto \omega_i$ and let B denote the product of the B_i. Define the operator $T : B \to B$ such that

$$T(f_1, \ldots, f_{n_g})_i(\gamma_g, \gamma_i, \gamma_{-ig}, \varepsilon_i) = \frac{1}{\phi+1}(\gamma_g + \gamma_i + \varepsilon_i) + \frac{\phi}{\phi+1}\frac{1}{n_g-1}\sum_{j\neq i}\mathrm{E}(f_j(\gamma_g, \gamma_j, \gamma_{-jg}, \varepsilon_j)).$$

A strategy profile $(f_1^*, \ldots, f_n^*)$ is a (not necessarily symmetric) Bayes-Nash equilibrium if and only if it is a fixed point of T. A straightforward calculation shows that T is a contraction mapping. At any point $(\gamma_g, \gamma_{1g}, \ldots, \gamma_{n_g g}, \varepsilon_i)$,

[69] Note that $\bar{\gamma}_g$ is different from γ_g, which is the direct group marginal utility contribution for an individual's choice.

$$
\begin{aligned}
|(Tf)_i - (Tg)_i| &= \frac{\phi}{\phi+1}\frac{1}{n_g-1}\Big|\sum_{j\neq i}\mathrm{E}\Big(f_j(\gamma_g,\gamma_j,\gamma_{-jg},\varepsilon_j)\Big) \\
&\quad - \mathrm{E}(g_j(\gamma_g,\gamma_j,\gamma_{-jg},\varepsilon_j))\Big| \\
&\leq \frac{\phi}{\phi+1}||f-g||_\infty.
\end{aligned}
$$

Since T is a contraction, it has a unique fixed point, and so equilibrium is unique. □

A2. PROOF OF THEOREMS 3, 4, 5 AND 7 ON SOCIAL NETWORKS

For exclusive averaging Bramoullé, Djebbari, and Fortin (2009) have already proven that if the network is the union of groups, then I, A and A^2 are linearly dependent if and only if groups are all the same size. They also have shown that if the network is transitive and contains no groups, then $A^2 = 0$. All we need to show is that linear dependence implies transitivity, and that transitivity implies that the network is the union of weakly connected components each of which either has $A^2 = 0$ or is a group. For inclusive averaging we simply replicate the entire program.

We begin with an elaboration of theorem 4 which does not depend on how the weighted adjacency matrix is assembled. The proof of this theorem, when combined with theorem 2, implies theorem 4 in the text. Theorem 2 states that the failure of identification implies that I, A and A^2 are linearly dependent. Theorem A.2 states that if these matrices are linearly dependent, then the network must be transitive. If the network is both transitive and undirected, it must consist of the union of groups. It follows from theorem 2.i that if the hypothesis of theorem 4 is true, there are scalars λ_0, λ_1, and λ_2, not all 0, such that $\lambda_0 I + \lambda_1 A + \lambda_2 A^2 = 0$.

Theorem A.2. Characterization of networks satisfying the dependence hypothesis. *Let (V, E) be a network with a weighted adjacency matrix A such that $(I - JA)$ is invertible for all values of J. Suppose that $\lambda_0 I + \lambda_1 A + \lambda_2 A^2 = 0$ for some λ weights not all zero.*

i. *The network is transitive.*

ii. *If $\lambda_2 = 0$, the network is totally disconnected.*

iii. *If $\lambda_2 \neq 0$ and $\lambda_1 = 0$, with exclusive averaging, the network is the union of groups of size 2 or, if $A^2 = 0$, the network is bipartite directed. With inclusive averaging, $A = I$ and the network is totally disconnected.*

Proof of theorem A.2. If $\lambda_2 = 0$, then $\lambda_0 I + \lambda_1 A = 0$. If $\lambda_1 = 0$, then $\lambda_0 = 0$, contradicting the hypotheses of the theorem. Thus if $\lambda_2 = 0$, then $\lambda_1 \neq 0$ and the matrix A is diagonal. With exclusive averaging, $\lambda_0 = 0$ and $A = 0$; with inclusive averaging, $\lambda_0 = 1$ and $A = I$. In either case, the network is totally disconnected (and, in particular, transitive).

If $\lambda_2 \neq 0$, and $\lambda_1 = 0$, there is a scalar γ_0 such that $A^2 = \gamma_0 I$. With inclusive averaging, the row sums of A and A^2 are both 1, and so $\gamma_0 = 1$ and $A^2 = I$. If $(i, k) \in E$ and $(k, j) \in E$, then

$j = i$. If not, $[A^2]_{ij} > 0$ for $j \neq i$. Since A is non-negative, $A^2 = I$ implies that if $i \neq k$ and $a_{ij} > 0$, then $a_{jk} = 0$. Since $a_{kk} > 0$ for all k, it follows that for $i \neq k$, $a_{ik} = 0$. Thus $a_{ii} = 1$, and so $A = I$.

For exclusive averaging, suppose first that $A^2 \neq 0$. Then $\gamma_0 \neq 0$, and so A^2 is a diagonal matrix in which $[A^2]_{ii} = \gamma_0 > 0$. If $a_{ij} > 0$ and $a_{jk} > 0$, then $[A^2]_{ik} > 0$, so $k = i$. That is, there are no chains connecting three nodes, and so trivially, the network is transitive. Now $[A^2]_{ii} > 0$ if there is a j such that a_{ij} and a_{ji} are both positive. There cannot be another $k \neq j$ such that $a_{ik} > 0$, because then $[A^2]_{jk} > 0$. Thus the social network is a collection of marriages; groups of size 2. The other possibility is that $A^2 = 0$, which happens if and only if the network is either a directed bipartite graph ($A \neq 0$) or totally disconnected ($A = 0$). These graphs are transitive.

If $\lambda_2 \neq 0$ and $\lambda_1 \neq 0$, then $\gamma_1 \neq 0$ and for nodes $i \neq j$, $[A^2]_{ij} = \gamma_1 a_{ij}$. If there is path of length 2 from nodes i to j, then $[A^2]_{ij} > 0$, and hence $a_{ij} > 0$, so $(i, j) \in E$ and the network is transitive. □

We now turn to the proof of theorem 3. The remainder of this appendix explores the case where $\lambda_2 \neq 0$ and $\lambda_1 \neq 0$. Thus $A^2 = \gamma_1 A + \gamma_0 I$, with $\gamma_1 \neq 0$. To proceed we need some facts about transitive graphs.

Lemma A.1. *If (V, E) is transitive, then*

i. *If $(i, j) \in E$ for some $i \in V_g$ and $j \in V_h$, then for all $i' \in V_g$ and $j' \in V_h$, $(i', j') \in E$.*

ii. *The relation $V_g \succ V_h$ iff $(i, j) \in E$ for some $i \in V_g$ and $j \in V_h$ and $V_g \neq V_h$ is transitive and asymmetric.*

Each weakly connected component corresponds to a block of the blockdiagonal matrix A, and if the powers of A are linearly dependent then the powers of each diagonal block are too. So we begin by assuming the existence of only one block. The facts about transitive networks imply that the matrix A has the following structure:

$$A = \begin{pmatrix} A_{g_a,g_a} & A_{g_a,g_b} & \cdots & A_{g_a,g_c} \\ 0 & A_{g_b,g_b} & \cdots & A_{g_b,g_c} \\ 0 & 0 & & \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{g_c,g_c} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

Matrix A_{g_x,g_y} is of size $|V_{g_x}| \times |V_{g_y}|$. With exclusive averaging, each matrix on the diagonal has 0s on its diagonal, and is strictly positive off it. With inclusive averaging, diagonal entries are not 0. If a block A_{g_x,g_y} has only 0 blocks to its right, there must be a non-zero block above. Finally, for each i, j and k, $a_{ij} = a_{ik}$, and the row sums are 1.

Proof of theorem 3. First we take up the case of exclusive averaging.

Lemma A.2. *Assume exclusive averaging. $\gamma_0 \neq 0$ iff (V, E) is a group of size at least 2. If so, then*

$$A^2 = \frac{1}{|V|-1} I + \frac{|V|-2}{|V|-1} A. \tag{87}$$

Proof of lemma A.2. "If" is a calculation. For the other direction, suppose the two coefficients are positive. Then $[A^2]_{ii} = \gamma_0$ for all i since $a_{ii} = 0$. Suppose (V, E) is not a group. There must be a strongly connected component V_h which is minimal with respect to $\succ$, and another strongly connected component V_g such that $V_g \succ V_h$. Members of V_h connect only to themselves, and so the cardinality of V_h must exceed 1, or else $\gamma_0 = 0$. For $i \in V_g$ and $j \in V_h$,

$$[A^2]_{ii} = \frac{|V_g| - 1}{(|V_g| - 1 + |V_h| + m)^2}$$
$$[A^2]_{jj} = \frac{1}{|V_h| - 1}$$

where m is the number of nodes outside of V_h members of V_g are connected to. Both of these numbers must equal γ_0, and so $(|V_g| - 1)(|V_h| - 1) = (|V_g| - 1 + |V_h| - 1 + m + 1)^2$, which is impossible. Thus (V, E) is a group, and equation (87) follows. □

Now we identify the structure for the remaining case, which has $\gamma_1 > 0$ and $\gamma_0 = 0$. Note that directed bipartite networks are characterized by the property that $A^2 = 0$. Anyone who is influenced by someone influences no one.

Lemma A.3. *Assume exclusive averaging and that $\gamma_0 = 0$. Then $\gamma_1 = 0$, and the weakly connected component is either a singleton or a directed bipartite graph.*

Proof. From the assumption it follows that $A^n = \gamma_1^n A$ for all n. In particular, $[A^n]_{ii} = 0$ for all i and n. Thus there are no paths that return to their starting point, and so no path can have length more than $|V| - 1$. Thus $A^{|V|} = 0$ and it follows that $\gamma_1 = 0$ and $A^2 = 0$. If $A = 0$, then the weakly connected component has only one member. If $A \neq 0$, all paths are of length 1. The network is a directed bipartite graph. □

Now we repeat the same exercise for inclusive averaging. Again, assume (V, E) is weakly connected.

Lemma A.4. *Assume inclusive averaging. If (V, E) is a group, then $A^2 = A$, $\gamma_1 = 1$ and $\gamma_0 = 0$. If $A^2 = A$ and $A \neq 0$, then (V, E) is a group.*

Proof of Lemma A.4. If (V, E) is a group, then all elements of A are identical and the row sum is 1. Thus $A^2 = A$, and so forth.

Suppose now that (V, E) is not a group. Then there are subgraphs $(V_1, E_1), \ldots, (V_K, E_K)$ such that the V_k partition V each component is strongly connected, and maximally so. All edges not in $\cup_k E_k$ run between components. Transitivity implies that if there is an edge from any vertex in V_j to any vertex in V_k, then there is an edge from

every vertex in V_j to every vertex in V_k. Say there is a path from component (Vi, Ei) to component (V_j, E_j) if there is a path from any vertex in V_i to any vertex in V_j. Any such path must be acyclic. Otherwise, transitivity implies that all nodes in the components with all edges between them form a strongly connected component, contradicting the maximality of the (V_k, E_k). Thus there is at least one "terminal component", (V_1, E_1), from which there is no path to any other component. Since (V, E) is weakly connected and not a group, there must be a component (V_2, E_2) which connects directly to (V_1, E_1). Each element may be connected to vertices not in either V_1 or V_2. Choose $i \in V_2$, and let W denote the set of vertices not in $V_1 \cup V_2$ to which i is connected. Observe that

$$\begin{aligned}
[A^2]_{ii} &= \sum_{j\in W} \frac{1}{|V_1|+|V_2|+|W|} a_{ji} + \sum_{j\in V_2} \left(\frac{1}{|V_1|+|V_2|+|W|}\right)^2 \\
&\quad + \sum_{j\in V_1} \frac{1}{|V_1|+|V_2|+|W|} \frac{1}{|V_2|} \\
&= \frac{1}{|V_1|+|V_2|+|W|} \sum_{j\in W} a_{ji} + \frac{|V_2|}{(|V_1|+|V_2|+|W|)^2} \\
&\quad + \frac{|V_1|}{|V_1|+|V_2|+|W|} \frac{1}{|V_1|} \\
&= \frac{1}{|V_1|+|V_2|+|W|} + \frac{1}{|V_1|+|V_2|+|W|} \sum_{j\in W} a_{ji} + \frac{|V_2|}{(|V_1|+|V_2|+|W|)^2} \\
&> \frac{1}{|V_1|+|V_2|+|W|} = A_{ii}
\end{aligned}$$

since $|V_2| > 0$, so $A^2 \neq A$. □

Lemma A.5. *If $\lambda_2 \neq 0$, (V, E) is a group.*

Proof of lemma A.5. If (V, E) is not a group, each strongly connected component is of size 1. If not, and both i and j are in the same component, then

$$\gamma_0 + \gamma_1 a_{ii} = [A^2]_{ii} = [A^2]_{jj} = \gamma_1 a_{jj} = \gamma_1 a_{ii}. \tag{88}$$

Since $a_{ii} \neq 0$, $\gamma_0 = 0$ and $\gamma_1 = 1$ and so A is a group, which is impossible. If so, suppose that $\{j\}$ is minimal with respect to $\succ$ and that $\{i\} \succ \{j\}$. Let m denote the number of other nodes that influence i. Then

$$\begin{aligned}
[A^2]_{ii} &= \frac{1}{(2+m)^2} = \gamma_0 + \gamma_1 \frac{1}{2+m} \\
[A^2]_{ij} &= \frac{1}{(2+m)^2} + \frac{1}{2+m} = \gamma_0 + \gamma_1 \frac{1}{2+m}
\end{aligned}$$

and would imply $0 = 1/(2 + m)$ which is impossible. □

To complete the proof of theorem 3, consider first the case of exclusive averaging. If the parameters are not identified, then I, A and A^2 are dependent. In particular, the powers of each matrix block corresponding to a given weakly connected component are dependent. Each component is either a group, a singleton, or a directed bipartite network. Furthermore, Bramoullé *et al.* show that all groups must be the same size. Both singleton components and directed bipartite networks have individuals who are influenced by no one. It follows from Bramoullé *et al.* proposition 3 that if $k \neq 0$, the parameters are identified if such individuals exist. Consequently, all weakly connected components are groups of a common size, at least 2. Conversely, if all components are groups of the same size, there are no individuals influenced by no one, and identification fails according to lemma A.2 and theorem 2.ii.

In the case of inclusive averaging, it follows as before that the relevant matrix powers are linearly dependent. It follows from lemma A.5 that each component is either a group or a singleton. A singleton gives identification, so each component is a group of size at least 2. A calculation shows that the groups need not be the same size for the powers of the entire matrix to be linearly dependent. Conversely, if all components are groups, then identification fails according to lemma A.4 and theorem 2.ii. □

Proof of theorem 5. It suffices to prove the theorem for the open and dense set of matrices S_1 which are strictly positive. Then $\lambda_2 \neq 0$, so write $A^2 = \gamma_1 A + \gamma_0 I$. We need to prove the claim of the theorem for the set S_1 of matrices that can be written this way. The set S_1 is semi-algebraic and closed. It suffices to show that S/S_1 is dense.

Consider a matrix A in S, and denote its square by B. Consider matrices of the form $A(\varepsilon)$ whose i, j element is $a_{11} + \varepsilon$ for $i = j = 1$ and a_{ij} otherwise. Then

$$\begin{aligned} B(\varepsilon)_{11} &= b_{11} + 2a_{11}\varepsilon + \varepsilon^2, \\ B(\varepsilon)_{1j} &= b_{1j} + a_{1j}\varepsilon \quad \text{for } j \neq 1, \\ B(\varepsilon)_{i1} &= b_{i1} + a_{i1}\varepsilon \quad \text{for } i \neq 1, \\ B(\varepsilon)_{ij} &= b_{ij} \text{ otherwise.} \end{aligned}$$

Suppose that $B(\varepsilon)$ is in S for all small ε. Computing.

$$\begin{aligned} \gamma_1 &= (b_{21} + a_{21}\varepsilon)/a_{21} \\ \gamma_0 &= b_{22} - \gamma_1 a_{22}. \end{aligned}$$

Then the equation $B(\varepsilon)_{11} = \gamma_1 A(\varepsilon)_{11} + \gamma_0$ is a linear (not quadratic) equation in ε. A necessary condition for linear dependence of the powers of $A(\varepsilon)$ for more than 1 value of ε is that the coefficient on ε is zero. This happens only for a set of A-matrices S_2 of codimension at least 1. Hence for all but at most one small enough ε, $A(\varepsilon) \notin S_1/S_2$. Since S_2 is nowhere dense, this proves the theorem. □

Proof of theorem 7. Let $m = F(c, d, J, \gamma)$; M is the matrix of reduced form coefficients. Our goal is to see how they map back to the structural parameters. We will prove the theorem for n_v odd and equal to $2K + 1$. The proof for even n_v is similar.

By hypothesis, $I - JA(\gamma)$ is nonsingular. Thus

$$\begin{aligned} M &= (I - JA(\gamma))^{-1}(cI + dA(\gamma)) \\ &= (I - JA(\gamma))^{-1}(cI + dI - dI + dA(\gamma)) \\ &= (c + d)(I - JA(\gamma))^{-1} - dI \end{aligned} \tag{89}$$

and so

$$dI + M = (c + d)(I - JA(\gamma))^{-1}, \tag{90}$$

which verifies that $dI + M$ is nonsingular if $c + d \neq 0$.

In view of equation (57) which defines $A(\gamma)$,

$$(I - JA(\gamma))_{11} = 1$$

and

$$-\frac{(I - JA(\gamma))_{12}}{(I - JA(\gamma))_{11}} = J\gamma,$$

$$\frac{(I - JA(\gamma))_{13}}{(I - JA(\gamma))_{12}} = \cdots = -\frac{(I - JA(\gamma))_{1K+1}}{(I - JA(\gamma))_{1K}} = \gamma.$$

Now define

$$\begin{aligned} \mathcal{M} &= \{M : \text{for some } (c, d, J, \gamma) \in \mathcal{P}, F(c, d, J, \gamma) = M\} \\ \mathcal{M}_{dJ} &= \{M : \text{for some } (c, \gamma) \in \mathbf{R}^2 \times [0, 1), F(c, d, J, \gamma) = M\}. \end{aligned}$$

These are, respectively, the sets of all possible reduced form matrices and those reduced forms consistent with a particular parameter pair of structural parameters (d, J).

Equation (90) then requires the following: If $M \in \mathcal{M}_{dJ}$, then

$$(dI + M)_{11} \neq 0$$

and

$$-\frac{(dI + M)^{-1}_{12}}{(dI + M)^{-1}_{11}} = J\gamma, \tag{91}$$

$$\frac{(dI + M)^{-1}_{13}}{(dI + M)^{-1}_{12}} = \cdots = \frac{(dI + M)^{-1}_{1K+1}}{(dI + M)^{-1}_{1K}} = \gamma. \tag{92}$$

We will use this fact to show that for a given reduced form matrix $M \in \mathcal{M}$ there are at most $2(n_V - 1)$ possible values of (d, J) pairs consistent with equation (92). We will show that each of these (d, J) pairs is consistent with a unique (c, γ) pair, which proves the theorem.

Under our assumed model specification, M is symmetric. Thus it has real eigenvalues $\lambda_1, \ldots, \lambda_{n_v}$, and is diagonalizable by a unitary matrix P. Furthermore, for any

scalar d, $dI + M$ is diagonalized by the same matrix P, and has eigenvalues $d + \lambda_1, \ldots, d + \lambda_{n_v}$. Consequently,

$$(dI + M)^{-1} = P^{-1} \begin{pmatrix} \frac{1}{d+\lambda_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d+\lambda_2} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & \frac{1}{d+\lambda_{n_V}} \end{pmatrix} P.$$

The i, j'th entriy of the matrix product on the right is

$$(dI + M)^{-1}_{ij} = \frac{1}{\prod_j (d + \lambda_j)} \sum_k p_{ki} p_{kj} \prod_{l \neq k} (d + \lambda_l),$$

for any d which is not the negative of an eigenvalue of M (which would make $dI + M$ singular). We now ask, for which values of d can equation (92) be satisfied? Define $\phi_{ij}(d) = \sum_k p_{ik} p_{kj} \prod_{l \neq k} (d + \lambda_l)$. Equation (92) implies that

$$\frac{\phi_{13}(d)}{\phi_{12}(d)} = \frac{\phi_{14}(d)}{\phi_{13}(d)}$$

and so

$$p(d) \equiv \phi_{13}(d)^2 - \phi_{14}(d)\phi_{12}(d) = 0.$$

Then $p(d)$ is a polynomial of degree at most $2(n_V - 1)$. The dependence of the polynomial on the terms p_{ij} is the link between the reduced form coefficients and the structural parameters. To see that it is not identically 0, suppose that $M = F(c', d', J', \gamma')$. From equation (89), it follows that the value of the derivative of the matrix $dI + M$ at $d = d'$ is $(c' + d')^2 (I - J' A(\gamma')\,)^2$. From this fact, a calculation shows that if neither γ' nor J' are 0, and $c' + d' \neq 0$, then $p'(d') = 0$ only for the solutions to a polynomial equation in γ and J which is not identically 0. Thus $p'(d') \neq 0$ only on a lower-dimensional set $\mathcal{C}$ of (c', d', J', γ'). That is, for $M \in \mathcal{M}' = \{M : \text{for some } (c, d, J, \gamma) \in \mathcal{P}/\mathcal{C}, F(c, d, J, \gamma) = M\}$,

Off of this set, for any d sufficiently near to but not equal to d', $p(d') \neq 0$. Thus for $M \in \mathcal{M}' = \{M : \text{for some } (c, d, J, \gamma) \in \mathcal{P}/\mathcal{C}, F(c, d, J, \gamma) = M\}$, $p(d) = 0$ has at most 2 $(n_V - 1)$ solutions. For each d which is a root of $p(d)$, $-(dI + M)^{-1}_{12} / (dI + M)^{-1}_{11} = J_{yd}$ (equation (91)). The ratio of any other pair of adjacent entries of the $(dI + M)^{-1}$ matrix determines γ_d. Finally, c_d solves $(I - JA(\gamma_d)\,)(dI + M) - dI = cI$ for c. Suppose, then, that $M = F(c', d', J', \gamma')$ for $(c', d', J', \gamma') \in \mathcal{P}/\mathcal{C}$. If a parameter vector $(c'', d'', J'', \gamma'')$ is not such that d'' is a root of $p(d)$, or that c'', J'', and γ'' do not equal the corresponding c_d, J_d and γ_d, then $(c'', d'', J'', \gamma'')$ is not observationally equivalent to

(c', d', J', γ'). If $\gamma = 0$, then neither d nor J can be identified. In this case $M = cI$. Conversely, if $M = c'I$, from (90) either $J = 0$ and $d = 0$ or $\gamma = 0$. From equation (58), if $J = 0$ and $M = c'I$, then either $\gamma = 0$ or $d = 0$ as well, and $c' = c$. □

A3. EQUILIBRIUM PROPERTIES OF DISCRETE CHOICE MODELS WITH SOCIAL INTERACTIONS

This appendix describes some aspects of the group-level equilibria for discrete choice models of social interactions. The models we discuss are similar in structure to the quantal response equilibria first developed by McKelvey and Palfrey (1995). The social interactions and quantal response equilibria literatures have evolved independently; as is true for other cases of parallel development that we have noted, each literature would benefit from integration with the other.

i. Basic structure of the binary choice model with social interactions

We first outline the theoretical properties of the binary choice model with social interactions for a single group g, following Brock and Durlauf (2001a). As in the text, choices are coded so that $\omega_i \in \{-1, 1\}$. Define $h_i = k + cx_i + dy_g$. From the perspective of the equilibrium of the group, contextual effects act in a way analogous to a constant term, an observation that is used in the proof of theorem 4 on identification. The utility function for a given choice is

$$V_i(\omega_{ig}) = h_i\omega_{ig} - \frac{J}{2}\mathrm{E}((\omega_{ig} - \bar{\omega}_{-ig})^2) + \eta_i(\omega_{ig}) \tag{93}$$

where $\bar{\omega}_{-ig} = (I-1)^{-1}\sum_{j\neq i}\omega_{jg}$ and $\eta_i(\omega_{ig})$ is a choice-specific random utility term. In parallel to the linear in means model, there is a penalty for expected square deviations of i's choice from the mean choices of others. Since $\omega_{ig}^2 \equiv 1$,

$$-\frac{J}{2}(\omega_{ig} - \bar{\omega}_{-ig})^2 = J\omega_{ig}\bar{\omega}_{-ig} - \frac{J}{2}(1 + \bar{\omega}_{-ig}^2).$$

The second term on the right is independent of ω_{ig}, and so the utility function of equation (93) yields the same behaviors as

$$V_i(\omega_{ig}) = h_i\omega_{ig} + J\omega_{ig}m_{ig}^e + \eta_i(\omega_{ig})$$

where $m_{ig}^e = (I-1)^{-1}\sum_{j\neq i}E\omega_{ig}|F_i$. It is immediate that

$$V_i(1) - V_i(-1) = 2h_i + 2Jm_{ig}^e - \varepsilon_i \tag{94}$$

where $\varepsilon_i = \eta_i(-1) - \eta_i(1)$. This justifies equation (59). As the group size grows large, m_{ig}^e will become independent of i.

ii. Equilibria under logit models of social interactions

To illustrate the qualitative properties of the binary choice model with social interactions, following Brock and Durlauf (2001a), we maintain the i.i.d. error assumptions (60) and (61) and further assume a functional form for F_ε:

$$F_\varepsilon(z) = \frac{1}{1 + \exp(-\beta z)}$$

that is, the individual-specific utility terms errors are negative exponentially distributed. The parameter β indexes the degree of unobserved heterogeneity. A larger β implies less heterogeneity in the sense that the probability mass of $F_\varepsilon(z)$ is more concentrated towards the origin.

This functional form, when combined with equation (94), produces the canonical logistic density for equilibrium choices

$$\mu(\omega_{ig}|h_i, m_g) = \frac{\exp(\beta h_i \omega_{ig} + \beta J m_g \omega_{ig})}{\exp(\beta h_i \omega_{ig} + \beta J m_g \omega_{ig}) + \exp(-\beta h_i \omega_{ig} - \beta J m_g \omega_{ig})}. \tag{95}$$

From equation (95) it is immediate that the expected value of agent i's choice is

$$\begin{aligned} \mathrm{E}(\omega_{ig}|h_i, m_g) &= \frac{\exp(\beta h_i \omega_{ig} + \beta J m_g \omega_{ig}) - \exp(-\beta h_i \omega_{ig} - \beta J m_g \omega_{ig})}{\exp(\beta h_i \omega_{ig} + \beta J m_g \omega_{ig}) + \exp(-\beta h_i \omega_{ig} - \beta J m_g \omega_{ig})} \\ &= \tanh(\beta h_i + \beta J m_g). \end{aligned} \tag{96}$$

The expected group mean is simply the unweighted average of (96) across i. Letting $dF_{h|g}$ denote the empirical density of h_i within group g, m_g is implicitly defined by

$$m_g = \int \tanh(\beta h + \beta J m_g) dF_{h|g}.$$

To understand the properties of the equilibrium, we consider the baseline case in the literature in which h_i is constant, that is, for all i, $h_i \equiv h$, so that the equilibrium expected average choice levels are described by a functional equation.

$$m_g = \tanh(\beta h + \beta J m_g). \tag{97}$$

No closed form solution exists. Brock and Durlauf (2001a) characterize the properties of solutions to equation (97), which we summarize as

Theorem A.3. Equilibria in the logistic version of the binary choice model with social interactions.

i. *If* $\beta J < 1$, *equation* (97) *has a unique solution.*

ii. *If* $\beta J > 1$, *then there exists a nondecreasing and positive function* $\bar{h}(\beta J)$ *of* βJ *such that*

i. *the equilibrium solution to equation* (97) *is unique if* $|h| > \bar{h}\,(\beta J)$*; and*

ii. *there exist three equilibrium solutions to equation* (97) *if* $|h| < \bar{h}\,(\beta J)$*. One equilibrum has the same sign as h.*

The intuition for the theorem is straightforward. $\beta J < 1$ means that the endogenous social interaction effect is too weak to generate multiple equilibria. Notice that strength of the interaction effect is not determined by J, the endogenous effect parameter, in isolation, but is multiplied by the measure for heterogeneity. Why would a small value of β work against the existence of multiple equilibria? A small β implies fatter tails for the unobserved heterogeneity density. By symmetry of this density, fat tails means a relatively large fraction of the population will, in expectation, have their choices determined by their heterogeneity draws. This leaves too small a fraction whose behavior can exhibit multiple equilibria via self-consistent bunching; the utility differential between the choices is insufficiently affected by the range of possible m_g values once the tail draws are accounted for. In contrast, $\beta J > 1$ means that the endogenous social utility payoff is large enough relative to the symmetrically distributed heterogeneity, then multiple expected average choice levels are possible, if $|h| < \bar{h}\,(\beta J)$. Why is this second condition needed? If the common private incentive h has sufficient magnitude, it will determine a sufficiently large fraction of choices so that self-consistent bunching is not possible. Again, greater heterogeneity reinforces this effect. Notice that qualitative changes in the number of equilibria for this model occur in neighborhoods of the value 1 for βJ and $\bar{h}\,(\beta J)$ for h. These are bifurcation thresholds.

Blume and Durlauf (2003) extend this theorem by considering a dynamic analog of the binary choice model with social interactions. Their analysis focuses on the stability of the rational expectations equilibria associated with (97). For a dynamic analog of the model we have outlined, one can show the population spends most of its time in the vicinity of the equilibrium that maximizes average utility in the group, which is the equilibrium whose mean choice has the same sign as h.

iii. Generalizations of the binary choice model

The properties of this model generalize to a number of interesting related structures. For example, one can analyze the general preference specification

$$V_i(1) - V_i(-1) = h_i + Jm_g - \beta^{-1}\varepsilon_i$$

where F_ε is is an arbitrary probability distribution function for the unobservable individual-level ε heterogeneity. Retaining the i.i.d. error assumptions (60) and (61) and focusing on the case where h is constant, Brock and Durlauf (2006) prove that a close analog to theorem A.3 holds for this general binary choice model. Two changes occur when the logit function form assumption is dropped. First, the necessary condition for

multiple equilibria takes the form $\beta J > T$, where the threshold T cannot be determined without specification of F_ε. In other words, some threshold for βJ always exists that can produce multiple equilibria. Second, for part ii.b, the threshold result for multiple equilibria states that at least three equilibria exist. The more precise structure of theorem A.3 derives from the specific functional form found in equation (97). The qualitative features of the theorem do not.

The qualitative properties of the theorem also extend to local interactions environments, i.e., contexts where individuals are arrayed in some social space and only interact with their suitably defined neighbors. One version of a local interactions model is studied in Blume (1993). An expectational version of his model can be represented by

$$V_i(1) - V_i(-1) = h_i + J \sum_{|i-j|=1} \mathrm{E}(\omega_{jg}) - \varepsilon_i,$$

where $\mathrm{E}(\omega_{jg})$ is the equilibrium expectation of ω_{jg} conditional on the values of h_i across the population. If we impose the assumption that for all i, $h_i = h$, then it is immediate that equation (97) continues to characterize the symmetric equilibrium average choice levels in the population. It is obvious that other interactions structures can do the same. The similar aggregate properties for different interactions specifications is itself known as the property of universality, which in social interactions contexts means that there exist dimensions along which the qualitative properties of the models do not depend on the details of the interaction structure. The reader should consult Ioannides (2006) for extensions of these types of models to more complex interactions structures.

iv. Multinomial choice models with social interactions

Multiple equilibria and bifurcations are not unique to the binary choice context. Brock and Durlauf (2006) show that theorem A.3 is a special case of

Theorem A.4. Multiple equilibria in the multinomial logit model with social interactions. *Suppose that individual choices are characterized by equations* (71), (72), *and* (73). *Assume that* $h_{il} = k$ *for all* i *and* l. *If* $\beta J > L$, *there will exist at least three self-consistent choice probabilities*.

The dependence of the threshold on the number of choices L is intuitive. The larger the number of choices, under independence of ε_{il} across l, the greater the probability that one of the draws will dominate the agent's choice, which reduces the fraction of agents whose behavior can exhibit self-consistent bunching. Brock and Durlauf (2006) additionally provide analogous results for general density functions for ε_{il}. As in the binary choice case the results are less precise. This theorem and its generalization in Brock and Durlauf (2006) explain various simulation results in Bayer and Timmins (2005).

REFERENCES

Aizer, A., Currie, J., 2004. Networks or Neighborhoods? Interpreting Correlations in the Use of Publicly-Funded Maternity Care in California. J. Public Econ. 88, 2573–2585.

Akerlof, G., 1997. Social Distance and Social Decisions. Econometrica 65, 1005–1027.

Akerlof, G., Kranton, R., 2000. Economics and Identity. Q. J. Econ. 115, 715–753.

Akerlof, G., Kranton, R., 2002. Identity and Schooling: Some Lessons for the Economics of Education. J. Econ. Lit. 40, 1167–1201.

Altonji, J., Huang, C., Taber, C., 2005. Estimating the Cream Skimming Effect of Private School Vouchers on Public School Students. Unpublished, University of Wisconsin.

Amaro de Matos, J., Perez, J., 1991. Fluctuations in the Curie-Weiss Version of the Random Field Ising Model. J. Stat. Phys. 62, 587–608.

Anderson, S., de Palma, A., Thisse, J., 1992. Discrete choice theory of product differentiation. MIT Press, Cambridge.

Angrist, J., Lang, K., 2004. Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. Am. Econ. Rev. 94, 1613–1634.

Anselin, L., 2010. Thirty Years of Spatial Econometrics. Pap. Reg. Sci. 89 (3–25).

Aradillas-Lopez, A., 2008. Using Rationalizable Bounds for Conditional Choice Probabilities as Control Variables. Unpublished, University of Wisconsin, Madison.

Aradillas-Lopez, A., 2009. Nonparametric Probability Bounds for Rationalizable Choices in Strategic-Interaction Models with Approximate Equilibrium Beliefs. University of Wisconsin, Madison.

Aradillas-Lopez, A., 2010. Semiparametric Estimation of a Simultaneous Game with Incomplete Information. J. Econom. 157, 409–431.

Aradillas-Lopez, A., Tamer, E., 2008. The Identification Power of Equilibrium in Simple Games. J. Bus. Econ. Stat. 26, 261–283.

Arcidiacono, P., Foster, G., Goodpaster, N., Kinsler, J., 2009. Estimating Spillovers Using Panel Data, with an Application to the Classroom. Unpublished, Duke University.

Austen-Smith, D., Fryer Jr., R., 2005. An Economic Analysis of Acting White. Q. J. Econ. 120, 551–583.

Baier, A., 1986. Trust and Antitrust. Ethics 96, 231–260.

Bajari, P., Hong, H., Krainer, J., Nekipelov, D., forthcoming. Estimating Static Models of Strategic Interaction. J. Bus. Econ. Stat..

Bala, V., Goyal, S., 2000. A Noncooperative Model of Network Formation. Econometrica 68, 1181–1229.

Bass, F., 1969. A New Product Growth Model for Consumer Durables. Manag. Sci. 15, 215–227.

Bayer, P., Ferreira, F., McMillan, R., 2007. A Unified Framework for Measuring Preferences over Schools and Neighborhoods. J. Polit. Econ. 1125, 588–638.

Bayer, P., Ross, S., 2009. Identifying Individual and Group Effects in the Presence of Sorting: A Neighborhood Effects Approach. University of Connecticut.

Bayer, P., Ross, S., Topa, G., 2008. Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. J. Polit. Econ. 116, 1150–1196.

Bayer, P., Timmins, C., 2005. On the Equilibrium Properties of Locational Sorting Models. J. Urban Econ. 57, 462–477.

Bayer, P., Timmins, C., 2007. Estimating Equilibrium Models of Sorting Across Locations. Econ. J. 117, 353–374.

Becker, G., 1971. The economics of discrimination, second ed. University of Chicago Press, Chicago.

Becker, G., 1973. A Theory of Marriage: Part I. J. Polit. Econ. 81, 813–846.

Becker, G., Tomes, N., 1979. An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. J. Polit. Econ. 87, 1153–1189.

Ben Akiva, M., 1973. Structure of Passenger Travel Demand Models. Ph.D. Dissertation, Department of Civil Engineering, MIT.

Bénabou, R., 1993. Workings of a City: Location, Education, and Production. Q. J. Econ. 108, 619–652.

Bénabou, R., 1996. Equity and Efficiency in Human Capital Investment: The Local Connection. Rev. Econ. Stud. 63, 237–264.

Beresteanu, A., Molchanov, I., Molinari, F., 2008. Sharp Identification Regions in Games. Cornell University.
Berg, J., Dickhaut, J., McCabe, K., 1995. Trust, Reciprocity, and Social History. Games Econ. Behav. 10, 122–142.
Bernardo, J., Smith, A., 1994. Bayesian theory. John Wiley and Sons, New York.
Berry, S., Levinson, J., Pakes, A., 1995. Automobile Prices in Market Equilibrium. Econometrica 63, 841–890.
Berry, S., Tamer, E., 2006. Identification in Models of Oligopoly Entry. In: Blundell, R., Newey, W., Persson, T. (Eds.), Advances in economics and econometrics. Cambridge University Press, Cambridge.
Bertrand, M., Luttmer, E., Mullainathan, S., 2000. Network Interactions and Welfare Cultures. Q. J. Econ. 115, 1019–1055.
Binder, M., Pesaran, M., 2001. Life Cycle Consumption Under Social Interactions. J. Econ. Dyn. Control 25, 35–83.
Bisin, A., Horst, U., Özgür, O., 2006. Rational Expectations Equilibria of Economies with Local Interactions. J. Econ. Theory 74–116.
Blume, L., 1993. The Statistical Mechanics of Strategic Interaction. Games Econ. Behav. 5, 387–424.
Blume, L., Durlauf, S., 2001. The Interactions-Based Approach to Socioeconomic Behavior. In: Durlauf, S., Young, H. (Eds.), Social dynamics. MIT Press, Cambridge MA.
Blume, L., Durlauf, S., 2003. Equilibrium Concepts for Social Interaction Models. International Game Theory Review 5, 193–209.
Blume, L., Durlauf, S., 2005. Identifying Social Interactions: A Review. In: Oakes, J., Kaufman, J. (Eds.), Methods in social epidemiology. Jossey-Bass, San Francisco.
Bochnak, J., Coste, M., Roy, M, 1998. Real algebraic geometry. Springer-Verlag, Berlin.
Börsch-Supan, A., 1990. On the Compatibility of Nested Logit Models with Utility Maximization. J. Econom. 43, 373–388.
Boucher, V., Bramoullé, Y., Djebbari, H., Fortin, B., 2010. Do Peers Affect Student Achievement? Evidence from Canada Using Group Size Variation. CIRANO Working Paper 2010s-08.
Bramoullé, Y., Djebbari, H., Fortin, B., 2009. Identification of Peer Effects Through Social Networks. J. Econom. 150, 41–55.
Briggs, X., Popkin, S., Goering, J., 2010. Moving to Opportunity: The Story of an American Experiment to Fight Ghetto Poverty. Oxford University Press, Oxford.
Brock, W., 1993. Pathways to Randomness in the Economy: Emergent Nonlinearity and Chaos in Economics and Finance. Estud. Econ. 8, 3–55.
Brock, W., 2006. Tipping Points, Abrupt Opinion Changes, and Punctuated Policy Change. In: Repetto, R. (Ed.), Punctuated equilibrium and the dynamics of U.S. environmental policy. Yale University Press, New Haven.
Brock, W., Durlauf, S., 2001a. Discrete Choice with Social Interactions. Rev. Econ. Stud. 68, 235–260.
Brock, W., Durlauf, S., 2001b. Interaction-Based Models. In: Heckman, J., Leamer, E. (Eds.), Handbook of econometrics. North-Holland, Amsterdam.
Brock, W., Durlauf, S., 2001c. Growth Empirics and Reality. World Bank Econ. Rev. 15, 229–272.
Brock, W., Durlauf, S., 2002. A Multinomial Choice Model with Neighborhood Interactions. Am. Econ. Rev. 92, 298–303.
Brock, W., Durlauf, S., 2006. Multinomial Choice with Social Interactions. In: Blume, L., Durlauf, S. (Eds.), The economy as an evolving complex system. 3, Oxford University Press, Oxford.
Brock, W., Durlauf, S., 2007. Identification of Binary Choice Models with Social Interactions. J. Econom. 140, 52–57.
Brock, W., Durlauf, S., 2010. Adoption Curves and Social Interactions. J. Eur. Econ. Assoc. 8, 232–251.
Brock, W., Durlauf, S., West, K., 2003. Policy Analysis in Uncertain Economic Environments (with discussion). Brookings Pap. Econ. Act. 1, 235–322.
Bryk, A., Raudenbush, S., 2001. Hierarchical linear models. Sage Publications, Newbury Park.
Cabral, L., 1990. On the Adoption of Innovations with 'Network' Externalities. Math. Soc. Sci. 19, 299–308.

Calabrese, S., Epple, D., Romer, T., Sieg, H., 2006. Local Public Good Provision: Voting, Peer Effects, Mobility. J. Public Econ. 90, 959–981.
Calvó-Armengol, A., Patacchini, E., Zenou, Y., 2009. Peer Effects and Social Networks in Education. Rev. Econ. Stud. 76, 1239–1267.
Camerer, C., 2003. Behavioral game theory: Experiments in strategic interaction. Princeton University Press, Princeton.
Card, D., Mas, A., Rothstein, J., 2008. Tipping and the Dynamics of Segregation. Q. J. Econ. 123, 177–218.
Chamberlain, G., 1984. Panel Data. In: Griliches, Z., McFadden, D. (Eds.), Handbook of econometrics. 2, North Holland, Amsterdam.
Cipollone, P., Rosolia, A., 2007. Social Interactions in High School: Lessons from an Earthquake. Am. Econ. Rev. 97, 948–965.
Cohen-Cole, E., 2006. Multiple Groups Identification in the Linear-in-means Model. Econ. Lett. 92, 157–162.
Coleman, J., 1988. Social Capital in the Creation of Human Capital. Am. J. Sociol. 94, S95–S120.
Conley, T., Topa, G., 2002. Socio-Economic Distance and Spatial Patterns in Unemployment. Journal of Applied Econometrics 17, 303–327.
Conley, T., Topa, G., 2003. Identification of Local Interaction Models with Imperfect Location Data. Journal of Applied Econometrics 18, 605–618.
Conley, T., Topa, G., 2007. Estimating Dynamic Local Interactions Models. J. Econom. 140, 282–303.
Cooley, J., 2008. Desegregation and the Achievement Gap: Do Diverse Peers Help?. University of Wisconsin, Madison.
Cooley, J., 2009. Alternative Mechanisms of Peer Spillovers in the Classroom: Implications for Identification and Policy. University of Wisconsin.
Cooper, R., John, A., 1988. Coordinating Coordination Failures in Keynesian Models. Q. J. Econ. 103, 441–463.
Copic, J., Jackson, M., Kirman, A., 2009. Identifying Community Structures from Network Data via Maximum Likelihood Methods. B. E. Journal of Theoretical Economics 9 (1), Article 30.
Corcoran, M., Gordon, R., Laren, D., Solon, G., 1992. The Association Between Men's Economic Status and their Family and Community Origins. J. Hum. Resour 27, 575–601.
Costa, D., Kahn, M., 2007. Surviving Andersonville: The Benefits of Social Networks in POW Camps. Am. Econ. Rev. 97, 1467–1487.
Crandall, D., Cosley, D., Huttenlocher, D., Kleinberg, J., Suri, S., 2008. Feedback Effects between Similarity and Social Influence in Online Communities. In: Proc. 14th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining.
Cronk, L., 2007. The Influence of Cultural Framing on Play in the Trust game: A Maasai Example. Evol. Hum. Behav. 28, 352–358.
Cunha, F., Heckman, J., 2007. The Technology of Skill Formation. Am. Econ. Rev. 97, 31–47.
Daley, D., Gani, J., 1999. Epidemic modelling: An introduction. Cambridge University Press, Cambridge.
Datcher, L., 1982. Effects of Community and Family Background on Achievement. Rev. Econ. Stat. 64, 32–41.
Davezies, L., d'Haultfoeuille, X., Foughre, D., 2009. Identification of Peer Effects Using Group Size Variation. Econom. J. 397–413.
De Giorgi, G., Pellizari, M., Redaelli, S., 2010. Identification of Social Interactions through Partially Overlapping Peer Groups. Am. Econ. J. Appl. Econ. 2, 241–275.
De Graaff, T., Florax, R., Nijkamp, P., Reggiani, A., 2002. A General Misspecification Test for Spatial Regression Models: Dependence, Heterogeneity, and Nonlinearity. J. Reg. Sci. 41, 255–276.
De Paula, A., 2009. Inference in a Synchronization Game with Social Interactions. J. Econom. 148, 56–71.
De Paula, A., Tang, X., 2010. Inference of Signs of Interaction Effects in Simultaneous Games with Incomplete Information. Unpublished, University of Pennsylvania.
Drewianka, S., 2003. Estimating Social Interactions in Matching Markets: Externalities in Spousal Search. Rev. Econ. Stat. 85, 408–423.
Dubin, J., McFadden, D., 1984. An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. Econometrica 52, 345–362.

Dufwenberg, M., Gneezy, U., 2000. Measuring Beliefs in an Experimental Lost Wallet Game. Games Econ. Behav. 30 (163–82).
Durlauf, S., 1993. Nonergodic Economic Growth. Rev. Econ. Stud. 60, 349–366.
Durlauf, S., 1996a. A Theory of Persistent Income Inequality. J. Econ. Growth 1, 75–93.
Durlauf, S., 1996b. Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality. In: Barnett, W., Gandolfo, G., Hilinger, C. (Eds.), Dynamic disequilibrium modelling. Cambridge University Press, Cambridge.
Durlauf, S., 2002. On the Empirics of Social Capital. Econ. J. 112, 459–479.
Durlauf, S., 2004. Neighborhood Effects. In: Henderson, J., Thisse, J. (Eds.), Handbook of regional and urban economics, vol. 4. North Holland, Amsterdam.
Durlauf, S., 2006. Groups, Social Influences, and Inequality: A Memberships Theory Perspective on Poverty Traps. In: Bowles, S., Durlauf, S., Hoff, K. (Eds.), Poverty Traps. Princeton University Press, Princeton.
Durlauf, S., Ioannides, Y., 2010. Social Interactions. Annual Review of Economics 2.
Durlauf, S., Seshadri, A., 2003. Is Assortative Matching Efficient? Econ. Theory 21, 475–493.
Durlauf, S., Tanaka, H., 2008. Understanding Regression Versus Variance Tests for Social Interactions. Econ. Inq. 46, 25–28.
Easterly, W., 2009. Empirics of Strategic Interdependence: The Case of the Racial Tipping Point. B. E. Journal of Macroeconomics 9, Article 25.
Ekeland, I., Heckman, J., Nesheim, L., 2004. Identification and Estimation of Hedonic Models. J. Polit. Econ. 112, S60–S109.
Ellis, R., 1985. Entropy, large deviations and statistical mechanics. Springer-Verlag, Heidelberg.
Epple, D., Romano, R., forthcoming. Peer Effects in Education: A Survey of Theory and Evidence. In: Benhabib, J., Bisin, A. (Eds.), Handbook of social economics. Elsevier, Amsterdam.
Epple, D., Sieg, H., 1999. Estimating Equilibrium Models of Local Jurisdictions. J. Polit. Econ. 100, 966–991.
Ermisch, J., Gambetta, D., 2006. People's Trust: The Design of a Survey-Based Experiment. Sociology Working Paper no. 2006-08, Oxford University.
Ermisch, J., Gambetta, D., Laurie, H., Siedler, T., Uhrig, S., 2009. Measuring People's Trust. J. R. Stat. Soc. Ser. A 172 (4), 749–769.
Evans, W., Oates, W., Schwab, R., 1992. Measuring Peer Group Effects: A Study of Teenage Behavior. J. Polit. Econ. 100, 966–991.
Fehr, E., Fischbach, U., von Rosenbladt, B., Schupp, J., Wagener, G., 2002. A Nation-wide Laboratory: Examining Trust and Trustworthiness by Integrating Behavioral Experiments into Representative Surveys. Schmollers Jahrbuch 122 (4), 519–542.
Fisher, F., 1959. Generalization of the Rank and Order Conditions for Identifiability. Econometrica 27 (3), 431–447.
Fisher, F., 1966. The identification problem in econometrics. McGraw-Hill, New York.
Galichon, A., Henry, M., 2008. Set Identification in Models with Multiple Equilibria. Unpublished, University of Montreal.
Gaviria, A., Raphael, S., 2001. School Based Peer Interactions and Juvenile Behavior. Rev. Econ. Stat. 83, 257–268.
Glaeser, E., Sacerdote, B., Scheinkman, J., 1996. Crime and Social Interactions. Q. J. Econ. 115, 811–846.
Glaeser, E., Sacerdote, B., Scheinkman, J., 2003. The Social Multiplier. J. Eur. Econ. Assoc. 1, 345–353.
Glaeser, E., Scheinkman, J., 2002. Non-Market Interactions. In: Dewatripont, M., Hansen, L., Turnovsky, S. (Eds.), Advances in Theory and Econometrics: Theory and Applications. Cambridge University Press, Cambridge.
Glaeser, E., Laibson, D., Scheinkman, J., Soutter, C., 2000. Measuring Trust. Q. J. Econ. 115 (3), 811–846.
Goering, J., Feins, J. (Eds.), 2003. Choosing a better life. Urban Institute Press, Washington DC.
Goldberger, A., 1991. A course in econometrics. Harvard University Press, Cambridge MA.
Graham, B., 2008. Identifying Social Interactions Through Conditional Variance Restrictions. Econometrica 76, 643–660.
Graham, B., 2010. Econometric Methods for the Analysis of Assignment Problems in the Presence of Complimentarity and Social Spillovers. In: Bisin, A., Benhabib, J., Jackson, M. (Eds.), Handbook of social economics. North Holland, Amsterdam.

Graham, B., Hahn, J., 2005. Identification and Estimation of the Linear-In-Means Model of Social Interactions. Econ. Lett. 88, 1–6.
Guryan, J., Kroft, K., Notowidigdo, M., 2009. Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments. Am. Econ. J. Appl. Econ. 1, 34–68.
Hansen, L., Sargent, T., 1980. Formulating and Estimating Dynamic Linear Rational Expectations Models. J. Econ. Dyn. Control 2, 7–46.
Hardin, R., 2002. Trust and trustworthiness. Russell Sage Foundation, New York.
Hausman, J., Taylor, W., 1981. Panel Data and Unobserved Individual Effects. Econometrica 49, 1377–1398.
Heckman, J., 1979. Sample Selection Bias as a Specification Error. Econometrica 47, 153–161.
Heckman, J., 1997. Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations. J. Hum. Resour. 32, 441–462.
Heckman, J., 1998. Detecting Discrimination. J. Econ. Perspect. 12, 101–116.
Heckman, J., 2000. Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective. Q. J. Econ. 115, 45–97.
Heckman, J., 2001. Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture. J. Polit. Econ. 109, 673–748.
Heckman, J., 2005. The Scientific Model of Causality. Sociol. Methodol. 35, 1–98.
Heckman, J., 2007. The Economics, Technology, and Neuroscience of Human Capability Formation. Proc. Natl. Acad. Sci. U.S.A. 104, 13250–13255.
Heckman, J., 2008. Econometric Causality. Int. Stat. Rev. 76, 1–27.
Heckman, J., Honoré, B., 1990. The Empirical Content of the Roy Model. Econometrica 58, 1121–1149.
Heckman, J., Urzua, S., Vytlacil, E., 2006. Understanding Instrumental Variables in Models with Essential Heterogeneity. Rev. Econ. Stat. 88, 389–432.
Heckman, J., Vytlacil, E., 2001. Policy-Relevant Treatment Effects. Am. Econ. Rev. 91, 107–111.
Henrich, J., et al., 2005. Economic Man in Cross-Cultural Perspective: Behavioral Experiments in 15 Small-Scale Societies. Behav. Brain Sci. 28 (6), 795–855.
Hoff, K., Sen, A., 2005. Homewnership, Community Interactions, and Segregation. Am. Econ. Rev. 95, 1167–1189.
Honoré, B., de Paula, A., 2010. Interdependent Durations. Rev. Econ. Stud. 77, 1138–1163.
Horowitz, J., 2009. Semiparametric and nonparametric methods in econometrics. Springer-Verlag, New York.
Horst, U., Scheinkman, J., 2006. Equilibria in Systems of Social Interactions. J. Econ. Theory 130, 44–77.
Hoxby, C., 2000. Peer Interactions in the Classroom: Learning from Gender and Race Variation. NBER Working Paper no 7867.
Hsiao, C., 1983. Identification. In: Griliches, Z., Intriligator, M. (Eds.), Handbook of econometrics. 1, North-Holland, Amsterdam.
Ioannides, Y., 2006. Topologies of Social Interactions. Econ. Theory 28, 559–584.
Ioannides, Y., 2008. Full Solution of an Endogenous Sorting model with Contextual and Income Effects. Unpublished, Tufts University.
Ioannides, Y., Loury, L., 2004. Job Information Networks, Neighborhood Effects, and Inequality. J. Econ. Lit. 42, 1056–1093.
Ioannides, Y., Zabel, J., 2008. Interactions, Neighborhood Selection and Housing Demand. J. Urban Econ. 63, 229–252.
Iyer, S., Weeks, M., 2009. Social Interactions, Ethnicity and Fertility in Kenya. Unpublished, Cambridge University.
Jackson, M., 2008. Social and economic networks. Princeton University Press, Princeton NJ.
Jackson, M., Wolinsky, A., 1996. A Strategic Model of Social and Economic Networks. J. Econ. Theory 71, 44–74.
Jovanovic, B., 1987. Micro Shocks and Aggregate Risk. Q. J. Econ. 102, 395–410.
Kapoor, M., Kelejian, H., Prucha, I., 2007. Panel Data Models with Spatially Correlated Error Components. J. Econom. 140, 97–130.

Kelejian, H., Prucha, I., 2010. Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances. J. Econom. 157, 53–67.
Kling, J., Liebman, J., Katz, L., 2007. Experimental Analysis of Neighborhood Effects. Econometrica 75, 83–119.
Kling, J., Ludwig, J., Katz, L., 2005. Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment. Q. J. Econ. 120 (87–130).
Knack, S., Keefer, P., 1997. Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. Q. J. Econ. 112, 1252–1288.
Krauth, B., 2006a. Simulation-Based Estimation of Peer Effects. J. Econom. 133, 243–271.
Krauth, B., 2006b. Social Interactions in Small Groups. Can. J. Econ. 39, 414–433.
Lalive, R., Cattaneo, M., 2009. Social Interactions and Schooling Decisions. Rev. Econ. Stat. 91, 457–477.
Lee, L., 1983. Generalized Econometric Models with Selectivity. Econometrica 51, 507–512.
Lee, L., 2007. Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors and Fixed Effects. J. Econom. 140, 333–374.
Lee, L., Liu, X., Lin, X., 2010. Specification and Estimation of Social Interaction Models with Network Structure, Contextual Factors, Correlation and Fixed Effects. Econom. J. 13 (2), 145–176.
Levine, D., 1998. Modeling Altrusim and Spitefulness in Experiments. Rev. Econ. Dyn. 1, 593–622.
Li, J., Lee, L., 2009. Binary Choice under Social Interactions: An Empirical Study With and Without Subjective Data on Expectations. Journal of Applied Econometrics 140, 333–374.
Lin, X., forthcoming. Identifying Peer Effects in Student Academic Achievement by Spatial Autoregressive Models with Group Unobservables. J. Labor Econ.
Lindsay, B., 1995. Mixture models: Theory, geometry and applications. Institute of Mathematical Statistics, Hayward.
Loury, G., 1981. Intergenerational Transfers and the Distribution of Earnings. Econometrica 49, 843–867.
Loury, G., 2002. The anatomy of racial inequality. Harvard University Press, Cambridge MA.
Ludwig, J., Duncan, G., Hirschfield, P., 2001. Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-Mobility Experiment. Q. J. Econ. 116, 655–680.
Manski, C., 1988. Identification of Binary Response Models. J. Am. Stat. Assoc. 83, 729–738.
Manski, C., 1993. Identification of Endogenous Social Effects: The Reflection Problem. Rev. Econ. Stud. 60, 531–542.
Manski, C., 2002. Identification of Decision Rules in Experiments on Simple Games of Proposal and Response. Eur. Econ. Rev. 46, 880–891.
Manski, C., 2003. Partial identification of probability distributions. Springer-Verlag, Heidelberg.
Manski, C., 2010. Identification of Treatment Response with Social Interactions. Unpublished, Northwestern University.
Mas, A., Moretti, E., 2009. Peers at Work. Am. Econ. Rev. 99, 112–145.
McFadden, D., 1974. Conditional Logit Analysis of Qualitative Choice Behavior. In: Zarembka, P. (Ed.), Frontiers in economics. Academic Press, New York.
McFadden, D., 1978. Modelling the Choice of Residential Location. In: Karlqvist, A., Lundqvist, L., Snickars, F., Weibull, J. (Eds.), Spatial interaction theory and planning models. North-Holland, Amsterdam.
McKelvey, R., Palfrey, T., 1995. Quantal Response Equilibria for Normal Form Games. Games Econ. Behav. 10, 6–38.
McManus, D., 1992. How Common is Identification in Parametric Models? J. Econom. 53, 5–23.
Milgrom, P., Roberts, J., 1990. Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities. Econometrica 58, 1255–1277.
Moffitt, R., 2001. Policy Interventions, Low-Level Equilibria, and Social Interactions. In: Durlauf, S., Young, H. (Eds.), Social dynamics. MIT Press, Cambridge.
Monderer, D., Shapley, L., 1996. Potential Games. Games Econ. Behav. 14, 124–143.
Navarro, S., 2008. Control Functions. In: Durlauf, S., Blume, L. (Eds.), The New Palgrave Dictionary of Economics. second ed. Palgrave MacMillan, Basingstoke.
Nesheim, L., 2002. Equilibrium Sorting of Heterogeneous Consumers across Locations: Theory and Empirical Implications. Unpublished CeMMAP Working Paper CWP08/02.

Oreopoulos, P., 2003. The Long-Run Consequences of Living in a Poor Neighborhood. Q. J. Econ. 118, 1533–1575.
Page, M., Solon, G., 2003a. Correlations between Sisters and Neighbouring Girls in Their Subsequent Income as Adults. Journal of Applied Econometrics 18, 545–562.
Page, M., Solon, G., 2003b. Correlations between Brothers and Neighbouring Boys in Their Adult Earnings: The Importance of Being Urban. J. Labor Econ. 21, 831–855.
Prat, A., 2002. Should a Team be Homogeneous. Eur. Econ. Rev. 46, 1187–1207.
Putnam, R., Leonardi, R., Nanetti, R., 1993. Making democracy work: Civic tradition in modern Italy. Princeton University Press, Princeton.
Raudenbush, S., Sampson, R., 1999. Ecometrics: Towards a Science of Assessing Ecological Settings, with Application to the Systematic Social Observation of Neighborhoods. Sociol. Methodol. 29, 1–41.
Repetto, R. (Ed.), 2006. Punctuated equilibrium and the dynamics of U.S. environmental policy. Yale University Press, New Haven.
Rosenbaum, J., 1995. Changing the Geography of Opportunity by Expanding Residential Choice: Lessons from the Gautreaux Program. Housing Policy Debate 6, 231–269.
Sacerdote, B., 2001. Peer Effects with Random Assignment: Results for Dartmouth Roommates. Q. J. Econ. 116, 681–704.
Sattinger, M., 1993. Assignment Models of the Distribution of Earnings. J. Econ. Lit. 31, 831–880.
Schelling, T., 1971. Dynamic Models of Segregation. J. Math. Sociol. 1, 143–186.
Searle, S., Casella, G., McCulloch, C., 2006. Variance components. John Wiley and Sons, New York.
Sirakaya, S., 2006. Recidivism and and Social Interactions. J. Am. Stat. Assoc. 101, 863–875.
Soetevent, A., Kooreman, P., 2007. A Discrete-Choice Model with Social Interactions: with an Application to High School Teen Behavior. Journal of Applied Econometrics 22, 599–624.
Solon, G., Page, M., Duncan, G., 2000. Correlations between Neighboring Children in their Subsequent Educational Attainment. Rev. Econ. Stat. 82, 383–392.
Streufert, P., 2002. The Effect of Underclass Social Isolation on School Choice. J. Public Econ. Theory 2, 461–482.
Tamer, E., 2003. Incomplete Simultaneous Discrete Response Model with Multiple Equilibria. Rev. Econ. Stud. 70, 147–167.
Tamer, E., 2008. Partial Linear Model. In: Blume, L., Durlauf, S. (Eds.), The New Palgrave Dictionary of Economics. second ed. Palgrave MacMillan, Basingstoke.
Topa, G., 2001. Social Interactions, Local Spillovers, and Unemployment. Rev. Econ. Stud. 68 (261–295).
Vives, X., 1990. Nash Equilibrium with Strategic Complementarities. Journal of Mathematical Economics 19, 305–321.
Wallis, K., 1980. Econometric Implications of the Rational Expectations Hypothesis. Econometrica 48, 49–73.
Weinberg, B., 2008. Social Interactions with Endogenous Associations. Unpublished, Ohio State University.
Weinberg, B., Reagan, P., Yankow, J., 2004. Do Neighborhoods Affect Hours Worked? Evidence from Longitudinal Data. J. Labor Econ. 22, 891–924.
West, K., 2008. Heteroskedasticity and Autocorrelation Corrections. In: Durlauf, S., Blume, L. (Eds.), The New Palgrave Dictionary of Economics. second ed. Palgrave MacMillan, Basingstoke.
White, H., 1980. A Heteroskedasticity Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica 48, 817–838.
Young, H.P., 2010. Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning. Am. Econ. Rev. 99, 1899–1924.
Young, H.P., Burke, M., 2001. Competition and Custom in Economic Contracts: A Case Study of Illinois Agriculture. Am. Econ. Rev. 91, 559–573.