ECONOMETRIC ANALYSIS OF MODELS WITH SOCIAL INTERACTIONS

BRENDAN KLINE AND ELIE TAMER

1. Introduction

Models with social interactions accommodate the possibility that an individual's outcome may be related to the outcomes, choices, treatments, and/or other characteristics of other individuals. Consequently, these models can involve simultaneity in the determination of the outcomes that does not exist in classical models of treatment response. Such models have been used to study alcohol use (e.g., Kremer and Levy (2008)), education outcomes (e.g., Epple and Romano (2011), Sacerdote (2011)), obesity (e.g., Christakis and Fowler (2007)), smoking (e.g., Powell, Tauras, and Ross (2005)), and substance use (e.g., Lundborg (2006)), among other applications.

Models of social interactions may often be written in terms of a response function $y_{ig}(\cdot)$ that relates the outcome y_{ig} of individual i in group g to the treatment d_{ig} of individual i in group g and the outcomes y_{-ig} and treatments d_{-ig} of other individuals in group g, according to $y_{ig} = y_{ig}(d_{ig}, d_{-ig}, y_{-ig})$. Factors other than the treatments and others' outcomes that affect the outcome are implicitly captured by the functional form of $y_{ig}(\cdot)$. In particular, if applicable, the network of interactions among the individuals in the group are implicitly captured by the functional form of $y_{ig}(\cdot)$. Moreover, as discussed further later in the chapter, the response function can be written to depend on different quantities. For example, the response function could be written in "reduced form" to depend only on the treatments but

University of Texas at Austin
Harvard University
E-mail addresses: brendan.kline@austin.utexas.edu, elietamer@fas.harvard.edu.

Date: Saturday 10th June, 2017. Some of this chapter had been previously distributed as "The empirical content of models with social interactions" and "Some interpretation of the linear-in-means model of social interactions" by the same authors.

not the outcomes of the other individuals in the group. In a model of infectious diseases, $y_{ig}(\cdot)$ might be the health outcome of individual i as a function of the vaccination status of individual i, the vaccination status of the other people in individual i's reference group, and health outcomes of the other people in individual i's reference group. Or for another example, in a model of classroom production, $y_{ig}(\cdot)$ might be the test score of student i as a function of the test scores of the students in student i's classroom. This chapter discusses a variety of specifications of these response functions, including the linear-in-means model and nonlinear models and links these response functions to models of social interactions in economics and the corresponding identification problems. In some settings, these response functions correspond to best response functions of an underlying game. In other settings, these response functions are statistical objects that are used to define a particular treatment effect. Both interpretations are considered.

Specifically, this chapter discusses two main issues relating to the econometrics of models of social interactions. First, this chapter discusses the identification of models of social interactions. Models with social interactions are typically estimated with data on many small groups. For each group, the data typically consist of the outcome and treatment of each individual and perhaps other data like the demographic characteristics. The identification question asks whether it is possible to use such data to recover information about the underlying model that generated the data. And second, this chapter discusses the interpretation and policy relevance of models of social interactions. An important consideration in the application of models of social interactions concerns the adequacy of the specification of the model. Models of social interactions inevitably involve assumptions, either explicit or implicit, about the nature of the interactions. These assumptions often times entail significant restrictions on behavior that are not necessarily implied by economic theory for all applications. These assumptions also concern the relationship between the observed data and the underlying model that generated the data. Therefore, such assumptions are important considerations in the assessment of the adequacy of models of social interactions for particular applications.

Another important consideration in the application of models of social interactions concerns policy relevance, which relates to the question of how to use the model to evaluate policy interventions.

The chapter leaves out other important issues, such as the nontrivial question of statistical inference. In addition, the chapter does not address the important problems related to social welfare, characterizing winners and losers in models with social interactions.

Finally, the literature on social interactions is large, and so this chapter cannot cover all important issues in the literature. Hence, this chapter is a complement rather than a substitute for other already existing reviews, perspectives, and works like Manski (2000), Glaeser and Scheinkman (2001), Scheinkman (2008), Durlauf and Ioannides (2010), Blume, Brock, Durlauf, and Ioannides (2011), Graham (2015), and de Paula (2016).

2. Identification of models of social interactions

2.1. **The linear-in-means model.**

The most commonly used model of social interactions is the linear-in-means model. There are many variants of the linear-in-means model, all of which are similar to the specification that is specifically considered in this section:

$$y_{ig} = \alpha + x_{ig}\beta + \left(\sum_{j\neq i} g^w_{ij} x_{jg}\right)\gamma + \left(\sum_{j\neq i} g^w_{ij} y_{jg}\right)\delta + \epsilon_{ig}, \tag{1}$$

where y_{ig} is the outcome of individual i in group g, x_{ig} are exogenous observables (i.e., "explanatory variables") relating to individual i in group g, and ϵ_{ig} are exogenous unobservables relating to individual i in group g. If $\gamma = 0$ and $\delta = 0$, then the linear-in-means model is a standard linear model for y_{ig}.

Social interactions are accommodated by including $\left(\sum_{j\neq i} g^w_{ij} x_{jg}\right)\gamma + \left(\sum_{j\neq i} g^w_{ij} y_{jg}\right)\delta$ in the model for y_{ig}. This allows that y_{ig} is affected by the explanatory variables of the other individuals in group g, and by the outcomes of the other individuals in group g. Specifically, the linear-in-means model assumes that explanatory variables of the other individuals in group g affect y_{ig} through the linear weighted sum $\sum_{j\neq i} g_{ij}^w x_{jg}$ and that outcomes of the other individuals in group g affect y_{ig} through the linear weighted sum $\sum_{j\neq i} g_{ij}^w y_{jg}$.

Therefore, g_{ij}^w is a measure of the weighted social influence that individual j in group g has on individual i in group g. These weighted social influences are derived from an underlying network of connections amongst individuals in group g, represented by an adjacency matrix g. By definition, $g_{ij} = 1$ means that individual j in group g is "connected to" individual i in group g, and therefore has influence on individual i in group g. Conversely, $g_{ij} = 0$ means that individual j in group g is "not connected to" individual i in group g, and therefore does not have an influence on individual i in group g. Typically, these connections are friendships or other social relations. This network is usually (but not always) assumed to be observed in the data, so that g_{ij} reflects whether individual i in group g has nominated individual j in group g as a social influence. Then, in the typical specification of the linear-in-means model,

$$g_{ij}^w = \begin{cases} 0 & \text{if } g_{ij} = 0 \\ \frac{1}{\sum_{j\neq i} g_{ij}} & \text{if } g_{ij} = 1 \end{cases}. \tag{2}$$

Consequently, in the typical specification of the linear-in-means model, $\sum_{j\neq i} g_{ij}^w x_{jg}$ is the average of the observable explanatory variables of the individuals that influence individual i in group g, and $\sum_{j\neq i} g_{ij}^w y_{jg}$ is the average of the outcomes of the individuals that influence individual i in group g. Therefore, as motivates the name of the model, the outcome y_{ig} depends linearly on the mean characteristics of the individuals that influence individual i in group g.

The analysis of the linear-in-means model proceeds by writing the model at the group level rather than the individual level, where group g has N_g individuals. Let Y_g be the $N_g \times 1$ vector that stacks the elements y_{ig}. Let X_g be the $N_g \times K$ vector that stacks the vectors x_{ig} for different individuals in different rows, where K is the dimension of the exogenous explanatory variables. Let ϵ_g be the $N_g \times 1$ vector that stacks the elements ϵ_{ig}. And let G^w be the $N_g \times N_g$ matrix of weighted social influence, with g^w_{ij} in row i and column j.

Then,

$$Y_g = 1_{N_g \times 1}\alpha + X_g\beta + G^w X_g \gamma + G^w Y_g \delta + \epsilon_g.$$

The standard linear-in-means model uses the assumption that X_g and G^w are mean independent of the unobservables, in the sense that $E(\epsilon_g | X_g, G^w) = 0$. The mean independence of G^w from ϵ_g requires that friendship or link formation is suitably independent of the unobservables.

The corresponding reduced form, which expresses the endogenous outcomes of all individuals in the group as a function of the exogenous variables of all individuals in the group, is

$$Y_g = (I - \delta G^w)^{-1}\left(1_{N_g \times 1}\alpha + X_g\beta + G^w X_g\gamma + \epsilon_g\right).$$

The existence of this reduced form depends on nonsingularity of $I - \delta G^w$. As long as $|\delta| < 1$, $I - \delta G^w$ is strictly diagonally dominant, and therefore nonsingular. Therefore, using the exogeneity assumption $E(\epsilon_g | X_g, G^w) = 0$, it follows that

$$E(Y_g | X_g, G^w) = (I - \delta G^w)^{-1}\left(1_{N_g \times 1}\alpha + X_g\beta + G^w X_g\gamma\right), \tag{3}$$

and therefore

$$E(y_{ig} | X_g, G^w) = e_i'(I - \delta G^w)^{-1}\left(1_{N_g \times 1}\alpha + X_g\beta + G^w X_g\gamma\right),$$

where e_i is the unit vector of length N_g, with 1 as the i-th element and 0 as every other element. By tracing out the effect of the exogenous explanatory variables on the outcomes, it is possible to recover $\frac{\partial E(y_{ig}|X_g, G^w)}{\partial X_g}$ from the data. Applying matrix calculus rules,

$$\frac{\partial E(y_{ig}|X_g, G^w)}{\partial X_g} = \beta e_i'(I - \delta G^w)^{-1} + \gamma e_i'(I - \delta G^w)^{-1} G^w. \tag{4}$$

Per the standard convention, the (k, j) element of $\frac{\partial E(y_{ig}|X_g,G^w)}{\partial X_g}$ is $\frac{\partial E(y_{ig}|X_g,G^w)}{\partial x_{jkg}}$, the marginal effect of the k-th explanatory variable of individual j on the outcome of individual i. Therefore, the effect of X_g on the outcome depends in a complicated way on all of the model parameters, β, γ, and δ, and also the weighted social influence matrix G^w. This reflects the fact that manipulating X_g has three interrelated effects on y_{ig}.

First, the exogenous explanatory variables of individual i have a direct effect on the outcome of individual i, per the parameter β. If $\gamma = 0$ and $\delta = 0$, then $\frac{\partial E(y_{ig}|X_g,G^w)}{\partial X_g} = \beta e_i'$ so this would be the entire effect of the exogenous explanatory variables, with only individual i's exogenous explanatory variables having an effect on the outcome of individual i. However, note that if $\gamma \neq 0$ and $\delta \neq 0$, then β does not fully summarize the effect of the exogenous explanatory variables of individual i on the outcome of individual i, because manipulating the exogenous explanatory variables of individual i will influence the outcomes of the other individuals in the group, which will influence the outcome of individual i. In other words, β describes the "effect" of the exogenous explanatory variables of individual i on the outcome of individual i "holding fixed" the outcomes of the other individuals, which generically would also change in response to manipulating the exogenous explanatory variables of individual i.

Second, the exogenous explanatory variables of the other individuals have a direct effect on the outcome of individual i, per the parameter γ and social influence matrix G^w. If $\delta = 0$, then $\frac{\partial E(y_{ig}|X_g,G^w)}{\partial X_g} = \beta e_i' + \gamma e_i' G^w$, so the exogenous explanatory variables of all individuals affect the outcomes of other individuals as determined by G^w. As above, note that if $\delta \neq 0$, then β and γ and G^w do not fully summarize the effect of the exogenous explanatory variables on the outcome of individual i, because manipulating the exogenous explanatory variables will influence the outcomes of the other individuals in the group, which will influence the outcome of individual i. In other words, β and γ and G^w describe the "effect" of the exogenous explanatory variables on the outcome of individual i "holding fixed" the outcomes of the other individuals, which generically would also change in response to manipulating the exogenous explanatory variables.

Third, the exogenous explanatory variables of all individuals have an indirect effect on the outcome of individual i through their effects on the outcomes, and then the effects of outcomes on other outcomes, according to the social influence matrix G^w and parameter δ.

In the data, only the combination of all of these effects of the exogenous explanatory variables is observed. This suggests that the model parameters may not be point identified under all conditions. Essentially, the question is whether the parameters β, γ, and δ can be recovered from either the reduced form in Equation 3 or the marginal effects in Equation 4.

Specifically, suppose that the underlying network is such that all individuals in the group have influence on all other individuals in the group, in the sense that $g_{ij} = 1$ for all $i \neq j$. This could arise in empirical applications particularly if group membership but not connections within a group are observed by the econometrician. Then,

$$(I - \delta G^w)^{-1} = u_1^{-1} I_{N_g \times N_g} - u_2 u_1^{-1} (u_1 + N_g u_2)^{-1} 1_{N_g \times N_g}$$

where $u_1 = 1 + \frac{\delta}{N_g - 1}$ and $u_2 = -\frac{\delta}{N_g - 1}$.

Then, the effect of the k-th explanatory variable of individual i on the outcome of individual i is the same as the effect of the k-th explanatory variable of individual j on the outcome of individual j, for all k and i, j. And, the effect of the k-th explanatory variable of individual j on the outcome of individual i is the same as the effect of the k-th explanatory variable of individual l on the outcome of individual i, for all k and $i \neq j$ and $i \neq l$. Therefore, there are a total of only $2K$ distinct effects of the explanatory variables, but $2K + 1$ parameters that determine the effects of the explanatory variables. Consequently, the model parameters should not be expected to be point identified from these marginal effects.

In assorted variants of the specification of the linear-in-means model discussed here, both Manski (1993) and Moffitt (2001) have established details of this source of non-identification in the linear-in-means model. Manski (1993) referred to this source of non-identification as the "reflection problem," based on the idea that it is difficult to distinguish between an individual's behavior and the behavior being "reflected" back on the individual.

The basic intuition for the identification analysis of the linear-in-means model comes from thinking about the linear-in-means model as a system of simultaneous equations. If all individuals influence all other individuals, then there are no exclusion restrictions in the structural form of the outcome y_{ig}, from Equation 1, in the sense that all exogenous explanatory variables of all individuals appear directly in the structural form of the outcome y_{ig}. Therefore, since that specification of the model does not have any exclusion restrictions, for similar reasons to identification analysis of general simultaneous equations model, the parameters from that specification of the model are not point identified (e.g., Wooldridge (2010, Chapter 9)).

Conversely, if not all individuals influence all other individuals, then there is an exclusion restriction in the structural form of the outcomes, and therefore the potential for valid instruments that could identify the parameters of the model. Specifically, suppose there is an individual l such that $g_{il} = 0$, so that individual l has no direct influence on individual i. And suppose further there is an individual j such that $g_{ij} = 1$ and $g_{jl} = 1$, so that individual j has a direct influence on individual i and individual l has a direct influence on individual j. Then, x_{lg} does not appear directly in the structural form of the outcome y_{ig}, but x_{lg} are relevant instruments for the outcome y_{jg} in the structural form of the outcome y_{ig}, since x_{lg} directly affects y_{lg}, which therefore directly affects y_{jg} since $g_{jl} = 1$, which therefore directly affects y_{ig} since $g_{ij} = 1$. This arrangement of the weighted social influences of individuals i, j, and l is known as an intransitive triad, and was the source of identification studied by Bramoullé, Djebbari, and Fortin (2009).

Specifically, Bramoullé, Djebbari, and Fortin (2009) formalize this identification strategy and establish that an important (nearly) sufficient condition for point identification of assorted variant specifications of the linear-in-means model is that I, G, and G^2 (and G^3 in some specifications) are linearly independent. If indeed there is the intransitive triad involving individuals i, j, and l, then it would follow that the (i,l) element of G^2 is strictly positive, since by applying the standard formula for matrix multiplication, that element of G^2 is weakly greater than $g_{ij}g_{jl}$, which is 1. Equivalently, since the (i,l) element of G^2 counts the number of paths of length two between individual i and l, it is at least 1, because of the path going through individual j. But the (i,l) element of I is 0 since $i \neq l$, and the (i,l) element of G is 0, since $g_{il} = 0$. Therefore, any linear combination of I and G would also have 0 as the (i,l) element, so G^2 cannot be a linear combination of I and G, and hence I, G, and G^2 are linearly independent. See also De Giorgi, Pellizzari, and Redaelli (2010).

An important consideration in the application of these identification strategies is whether or not this linear independence condition holds. It is a stylized fact that social networks tend to exhibit triadic closure, otherwise known as transitivity of links or clustering, as in "the friend of my friend is my friend." In other words, if two "sides" of a triad are present in a network, then the third "side" is likely to also be present. Therefore, although intransitive triads generically do exist in social networks, leading to the linear independence required above, they are less frequent than if links in the network formed independently without clustering. Another important consideration in the application of this identification strategy is whether or not there is sufficient intransitive triads to result in precise estimates.

Other sources of identification are also possible in the linear-in-means model. Lee (2007) and Davezies, d'Haultfoeuille, and Fougère (2009) show that variation in group size can result in point identification, in particular even if all individuals within a group influence all other individuals within that group, so that there are no intransitive triads as discussed above. The intuition for why variation in group size can result in point identification is evident in the marginal effects from the linear-in-means model in Equation 4. Even if G involves connections between all individuals in the group, variation in group size implies that G^w would be different for groups of different size. Hence, the effect of the exogenous explanatory variables on outcomes is different in groups of different sizes, providing additional sources of identification of the model parameters from the marginal effects in different group sizes. For example, in an extreme case of groups of size one, the effect of the exogenous explanatory variables would not involve any of the social interaction parameters, providing identification of the other parameters. An important consideration in the application of this identification strategy is whether or not it is reasonable to assume that the same model parameters apply to all group sizes. Another important consideration in the application of this identification strategy is whether or not there is sufficient variation in group size to result in precise estimates. In particular, if all groups are relatively large, then "variation" in group size may have a small impact on the marginal effects of the explanatory variables, potentially resulting in imprecise estimates. Specifically, approaches that rely on variation in group size can suffer from weak identification and weak instruments problems, in the sense of the inability of standard asymptotic analysis to capture finite sample behavior. Blume, Brock, Durlauf, and Jayaraman (2015) show a connection between groups of equal size and the condition that I, G and G^2 are linearly dependent (stacking groups into one network), as discussed also in de Paula (2016). Therefore, identification can be weak unless there is enough heterogeneity in the number of individuals per group. This chapter leaves out these issues of statistical inference.

Besides identification analysis, it is useful to note that the linear-in-means model can be interpreted as the equilibrium of a certain game with quadratic utility functions. Specifically, consider the utility function for individual i in group g given by

$$u_{ig}(y_{ig}, y_{-ig}) = \theta_{ig} y_{ig} - \frac{(1-\delta) y_{ig}^2}{2} - \frac{\delta}{2}(y_{ig} - \zeta_{ig})^2,$$

with corresponding first order condition that implies the response function

$$y_{ig} = \theta_{ig} + \delta \zeta_{ig},$$

which is the linear-in-means model from Equation 1 with $\theta_{ig} = \alpha + x_{ig}\beta + \left(\sum_{j \neq i} g^w_{ij} x_{jg}\right)\gamma + \epsilon_{ig}$ and $\zeta_{ig} = \sum_{j \neq i} g^w_{ij} y_{jg}$. Hence, the linear-in-means model may be interpreted as a game with quadratic utility functions, where each unit of outcome provides the individual with "individual" utility θ_{ig}, where the "cost" of the outcome is $\frac{(1-\delta)y_{ig}^2}{2}$, and where individuals want to conform their outcome to the aggregate outcome ζ_{ig} according to the parameter $\frac{\delta}{2}$ when $\delta > 0$ and "non-conform" their outcome to the aggregate outcome ζ_{ig} according to the parameter $\frac{\delta}{2}$ when $\delta < 0$. Therefore, in the leading case of $\delta > 0$, perhaps as expected given Equation 1, the linear-in-means model involves a preference for conformity of outcomes, which results in a particular linear relationship between outcomes. Of course, positive affine transformations of this utility function would also result in the linear-in-means model. See also Blume, Brock, Durlauf, and Jayaraman (2015) for a similar analysis for incomplete information versions of the linear-in-means model.

Other details of the specification of the linear-in-means model are important considerations in empirical applications. For example, the econometrician may specify the functional form of G^w as a function of G in at least two leading ways. In the typical specification, as discussed above through Equation 2, the structural equation for y_{ig} in Equation 1 depends on the average outcome of the individuals that influence individual i. Therefore, for example, if the outcome is the number of cigarettes smoked, then individual i is affected by the average number of cigarettes smoked by the individuals that influence individual i. Consequently, an individual with 1 friend that smokes 10 cigarettes is influenced equally to an individual with 2 friends that each smoke 10 cigarettes. In an alternative specification, $g^w_{ij} = g_{ij}$, so that the structural equation for y_{ig} in Equation 1 depends on the summed outcome of the individuals that influence individual i. Consequently, an individual with 1 friend that smokes 10 cigarettes is less influenced than an individual with 2 friends that each smoke 10 cigarettes. Liu, Patacchini, and Zenou (2014) consider a variant of the linear-in-means model that includes both modes of interactions.

Another important consideration in the specification of the linear-in-means model is whether the social interaction happens through the actual realized outcomes or the expected outcomes. As specified above in the structural equation for y_{ig} in Equation 1, y_{ig} depends on the actual realized y_{jg}. Alternatively, y_{ig} could be given a structural equation that depends on the expected value of the outcomes of the individuals that influence individual i. Viewing the linear-in-means model as a game, this corresponds respectively to a game of complete information and a game of incomplete information. Games of incomplete information may be more appropriate for social interactions among relatively large groups of relatively anonymous individuals, where the condition that individuals have private information may be more realistic. Games of complete information may be more appropriate for social interactions among relatively small groups of individuals, where the condition that individuals do not have private information may be more realistic. Further, pure strategy Nash equilibria of a game of complete information have the feature that the equilibrium outcomes are also in ex-post equilibrium, in the sense that by definition the individuals would not want to change their outcome even after observing the outcomes of the other individuals, making it a possible approximation to a long-run equilibrium of a dynamic process. By contrast, equilibria of a game of incomplete information tend to not involve realizations of equilibrium outcomes that are in ex-post equilibrium, in these sense that generally individuals would want to change their outcome after observing the realized outcomes of the other individuals.

It may be desirable also to incorporate unobserved group fixed effects, which can allow that outcomes depend on some features of the group shared by all individuals in the group and unobserved by the econometrician. Graham and Hahn (2005) show how to analyze identification of the linear-in-means model with group fixed effects using panel data methods, viewing the "cross-sectional" dimension as the groups and the "time" dimension as individuals in the groups. Bramoullé, Djebbari, and Fortin (2009) require further conditions on G to accommodate such group fixed effects.

Finally, it may be desirable to allow heterogeneity in the social interaction effects, which can allow that different pairs of individuals have different influences on each other. Masten (2017) shows identification results for a system of simultaneous equations model, and applies these results to the linear-in-means model.

2.2. **Non-linear models.** The linear-in-means model hypothesizes a particular linear relationship between the outcome of an individual and the outcomes of the other individuals in the group. Non-linear models hypothesize a non-linear relationship instead.

One motivation for using non-linear models of social interactions is to reduce the dependence on the implicit assumptions imposed by the linear-in-means model. As discussed in more detail in Section 3, the linear-in-means model is an example of a model of social interactions that makes a variety of behavioral assumptions that do not necessarily apply to all empirical applications. The linear model of a conditional expectation can be viewed as an approximation to a wide class of possibly non-linear conditional expectations, indeed even if the outcome is binary as in the linear probability model. However, the linear-in-means model does not share those same approximation properties, because it also involves assumptions about the nature of the social interaction. These issues are particularly salient when the outcome is discrete rather than continuous. For example, as discussed in more detail in Section 3, social interactions with discrete outcomes should be generally expected to involve issues like multiple equilibria, which is not a feature of the linear-in-means model, suggesting that it may not be a good approximation to settings with discrete outcomes.

Often, non-linear models applicable to social interactions are based on game theory, often when the outcome of the social interaction is discrete or even binary, and begin with a specification of the utility function. In these models, the outcome is the actions taken by individuals in the game. Consider for example a model of a social interaction when the outcome is binary, with utility functions given by

$$u_{ig}(1, y_{-ig}) = \alpha + x_{ig}\beta + \left(\sum_{j \neq i} g_{ij}^{w} y_{jg}\right)\delta + \epsilon_{ig} \text{ and } u_{ig}(0, y_{-ig}) = 0.$$

Thus, individual i is normalized to get 0 utility from taking action 0, and gets utility from taking action 1 that is similar to the specification of the linear-in-means model. Importantly, the exogenous explanatory variables of the other individuals in the group do not affect the utility of individual i, because they are generally used as excluded variables in the identification strategies. However, it would be possible to allow *some* but not *all* exogenous explanatory variables of other individuals to affect the utility of individual i, or to allow some exogenous explanatory variables to affect the utility of all individuals, by including those factors in x_{ig}. If $\delta > 0$, then the utility individual i gets from taking action 1 increases as a function of $\sum_{j \neq i} g_{ij}^{w} y_{jg}$, so that for example the utility from smoking increases as a function of the smoking decisions of the other individuals in the group. Such models have been considered by Bresnahan and Reiss (1991), Berry (1992), Tamer (2003), Krauth (2006), Soetevent and Kooreman (2007), Ciliberto and Tamer (2009), Bajari, Hong, and Ryan (2010), Fox and Lazzati (2016), Kline (2016), and in particular Kline (2015) includes the focus on application to social interactions by including the network interaction structure into the utility function. Similar models, except with incomplete information, have been considered by Brock and Durlauf (2001, 2007), Aradillas-Lopez (2010, 2012), Bajari, Hong, Krainer, and Nekipelov (2010), Tang (2010), Xu (2011), Grieco (2014), Lee, Li, and Lin (2014), Wan and Xu (2014), and Lewbel and Tang (2015). See also for example Berry and Tamer (2006) or Berry and Reiss (2007). Note that many of these models do not explicitly allow for a network to affect the utility functions, in which case the assumption would be that all individuals within a group affect all other individuals within that group. Such models may still be directly applicable to social interactions, and such models are without loss of generality if the social interaction concerns pairs of individuals (e.g., Card and Giuliano (2013) for "best friend" pairs).

Models of social interactions based on game theory also require the specification of a solution concept, a behavioral assumption that describes the predicted outcome of the model as a mapping from the utility functions. Even strong behavioral assumptions like pure strategy Nash equilibrium results in multiple equilibria outcomes, in which case the model only predicts that the outcome of the social interaction lies within some set of possible equilibria, but does not predict the outcome uniquely. Using weaker behavioral assumptions like mixed strategy Nash equilibrium or rationalizability would only exacerbate the multiplicity of outcomes. See also Section 3 for more discussion of the possibility of multiple equilibria in social interactions.

Consider the basic intuition for the most common identification strategy for these models, when the social interaction involves the outcome of whether individuals smoke cigarettes. In general, such models could result in multiple potential equilibrium outcomes for any given specification of utility functions. For example, consider a social interaction among pairs of friends. For a particular specification of utility functions, it could be an equilibrium for neither individual to smoke cigarettes, if both individuals find it undesirable to smoke alone. For the same specification of utility functions, it could also be an equilibrium for both to smoke cigarettes, if both individuals find it desirable to smoke alongside a friend despite not finding it desirable to smoke alone.

Despite these complications, it is possible to point identify the utility parameters if there is at least one exogenous explanatory variable per individual that appears exclusively in that individual's utility function. If so, those explanatory variables can be used to exogenously drive the smoking decisions of all *other* individuals, relative to some specific individual i. In particular, by taking particularly extreme values of those explanatory variables, the smoking decisions of the other individuals can be taken to be essentially exogenously either to smoke or not to smoke. Equivalently, this overcomes the problems relating to multiple equilibria, because it results in dominant strategies to either smoke or not to smoke, as appropriate depending on whether the exogenous explanatory variables of the individuals other than i are taken to be extremely positive or to be extremely negative. By tracing out the effect of those "exogenous" smoking decisions on the smoking decision of the individual i, it is possible to identify the utility parameters for that individual i.

Hence, an important consideration in the application of such an identification strategy is whether or not the application actually does have such a large support explanatory variable. Identification based on large support in these games are subject to similar concerns as in other areas of econometrics. Under weaker assumptions, the model parameters may only be partially identified, as in Aradillas-Lopez and Tamer (2008), Ciliberto and Tamer (2009), Beresteanu, Molchanov, and Molinari (2011), Kline (2015), or Kline and Tamer (2016).

Note that these identification strategies do not rely on the same restrictions on the network as do the identification strategies for the linear-in-means model in Section 2.1. Instead, these identification strategies rely on exogenous explanatory variables that each affect a specific individual but are excluded from the utility functions of all other individuals. So the exclusion restriction used in the linear-in-means model comes from the structure of the network, whereas the exclusion restriction used in the models of games comes from more standard excluded regressors.

Most of the literature on social interactions has focused on models that ignore the timing of the outcomes subject to social interactions. However, some social interactions are mainly about the timing of taking an action. For example, married couples might prefer to coordinate the timing of their retirement decisions (e.g., Honoré and de Paula (2011)). de Paula (2009) and Honoré and de Paula (2010) study a duration model with social interactions. An intuition for an identification strategy for such models is that social interactions in timing should result in simultaneous timings (e.g., simultaneous retirements) that could not be explained otherwise.

2.3. **Response functions.** Manski (2013) and Lazzati (2015) study partial identification of the response function that expresses the outcome of an individual as a function of the treatments of all individuals in the individual's reference group. Thus, generically, the response function $y_{ig}(\cdot)$ relates the outcome of individual i in group g to the treatment d_{ig} of individual i in group g and the treatments d_{-ig} of other individuals in group g. Such models do not explicitly model the possibility that outcomes are affected by the outcomes of other individuals. However, under certain assumptions, these response functions could be interpreted as the "reduced form" that maps the treatments to the outcomes, with an un-modeled intervening stage in which outcomes affect other outcomes. For example, in the context of the linear-in-means model, this reduced form would correspond to Equation 3. However, as discussed in more detail in Section 3, such a reduced form does not exist in all models. For example, if the social interaction induces multiple equilibria, then it is not possible to associate a unique outcome to particular profiles of treatments, since depending on the equilibrium that is selected, different outcomes are possible.

Identification of the response function faces the fundamental problem of causal inference. An individual's response function is observed at the observed profile of treatments, but is not observed at any other counterfactual profile of treatments. Therefore, similar to identification of response functions in the abscence of any social interactions (e.g., Manski (1997), Manski and Pepper (2000, 2009), Okumura and Usui (2014)), shape restrictions help to tighten the identified bounds on the response functions. For example, it can be assumed that the outcome of each individual is a weakly monotone function of the treatment(s) of all individuals in the group, which tightens the identified bounds. Such bounds would allow the econometrician to conclude, for example, bounds on the probability that a randomly selected individual would smoke as a function of a specified (possibly counterfactual) arrangement of policy intervention treatments affecting the individual and each of the individual's friends.

Kline and Tamer (2012) study partial identification of the best responses in complete information binary games. These games involve the decision between two possible actions per individual. The best response function describes the utility maximizing decision of a particular individual as a function of any counterfactual specification of decisions of the other individuals. Thus, generically, the response function $y_{ig}(\cdot)$ relates the utility maximizing outcome of individual i in group g to the outcomes y_{-ig} of other individuals in group g. For example, the best response function could be an individual's utility maximizing decision of whether or not to smoke cigarettes, as a function of the smoking decisions of the individual's friends. Under a variety of behavioral assumptions, including Nash equilibrium assumptions and non-equilibrium assumptions like level-k rationality, interval bounds on these best response functions can be derived and therefore estimated based on the literature on interval identified parameters. Such bounds would allow the econometrician to conclude, for example, bounds on the probability that a randomly selected individual would smoke as a function of a specified (possibly counterfactual) arrangement of smoking decisions of the individual's friends. The identification result depends on the specific set of assumptions maintained, but the intuition for the identification strategy is always to ask what an observed set of decisions must necessarily imply about the utility function, and therefore about utility maximizing decisions. For example, suppose the game models the decision to smoke cigarettes, and suppose the econometrician assumes that there is a positive peer effect, in the sense that a peer smoking increases the utility from smoking. Suppose, for simplicity of exposition, that the "peer group" is a pair of friends. And suppose a particular individual is observed to smoke, despite the friend in the pair not smoking. What would happen if the friend in the pair started smoking? It is possible to conclude that individual would *also* smoke if the friend in the pair smoked, under appropriate behavioral assumptions and the maintained assumption of a positive peer effect, since if it was already utility maximizing to smoke with a friend that does not smoke, it would also be utility maximizing to smoke with a friend that does smoke.

2.4. **Causal impacts of treatments mediated through networks.** In standard models of treatment response in which treatments are assigned and outcomes are observed at the assigned treatments, it may be of interest to learn how much of the treatment effect is *mediated* through the network. For example, if some treatment is assigned at random to individuals within a group, then it is possible to learn the group treatment effect. But, how much of this effect is due to the network of connections within the group? This mediated treatment effect can be examined in the context of (endogenous) networks. The approach in this section adapts the counterfactual notation approach to studying networks that is more common in the study of randomized experiments and abstracts away from the economics of games, multiple equilibria, best response functions, and other similar issues discussed elsewhere in this chapter. This approach is built toward estimating average treatment effects or similar objects. This approach is complementary to the "economic" approach taken in Section 3, where interest is focused on best response functions directly related to utility functions in the context of an economic model.

Let the outcome for individual i be y_i, which is determined by a response function according to

$$y_i = f(d_i, G_i, \epsilon_i),$$

where f is the response function, d_i is the binary treatment of individual i, G_i is the network (or some network summary statistic) for individual i, and ϵ_i is an arbitrary unobservable for individual i. Also, network G_i is determined by a response function according to

$$G_i = m(d_i, \eta_i),$$

where η_i is an arbitrary unobservable for individual i.

The functions $f(\cdot)$ and $m(\cdot)$ are not best response functions in a game theoretic sense as in Section 3, since they encode a selection mechanism that fills in an equilibrium choice. For example, the network formation model may have multiple equilibria, but the function

$G = m(d, \eta)$ delivers a unique network G. The unobservable η therefore must involve the unobserved role of a selection mechanism. For example, if the selection mechanism is modeled as a coin flip, then η must capture the unobserved "outcome" of that coin flip. So, these functions are statistical relationships that can be used to define average treatment effects. Whether these average treatment effects or these response functions are useful is application specific.

In counterfactual notation, the "overall" treated outcome is

$$y_i(1) = f(1, m(1, \eta_i), \epsilon_i)$$

and the "overall" untreated outcome is

$$y_i(0) = f(0, m(0, \eta_i), \epsilon_i).$$

Similarly, define the counterfactuals that separately specify the treatment and network as

$$y_i(d, G) = f(d, G, \epsilon_i).$$

The subsequent analysis can be conditioned on predetermined variables, and it is assumed that (y, d, G) are observed. In this setup, in addition to the standard "overall" average treatment effect

$$ATE \equiv E(y_i(1) - y_i(0)) = E\left(f(1, m(1, \eta_i), \epsilon_i) - f(0, m(0, \eta_i), \epsilon_i)\right),$$

there is the direct treatment effect at some fixed network G_0,

$$DTE(G_0) \equiv E(y_i(1, G_0) - y_i(0, G_0)) = E\left(f(1, G_0, \epsilon_i) - f(0, G_0, \epsilon_i)\right)$$

and the indirect treatment effect at some fixed treatment d_0,

$$ITE(d_0) \equiv E(y_i(d_0, m(1, \eta_i)) - y_i(d_0, m(0, \eta_i))) = E\left(f(d_0, m(1, \eta_i), \epsilon_i) - f(d_0, m(0, \eta_i), \epsilon_i)\right).$$

The direct treatment effect is the "effect" of the treatment on the outcome, holding fixed the network at some fixed/predetermined G_0. The indirect treatment effect is the "effect" on the outcome in response to a change in the network in response to a change in the treatment, holding fixed the "direct" treatment received by the individual at some fixed d_0.

2.4.1. *Identification with random assignment of treatment.* When the treatment d is randomly assigned, in the sense that $(\epsilon, \eta) \perp d$, the "overall" average treatment effect can be point identified in the standard way:

$$ATE \equiv E(y(1) - y(0)) = E\left(f(1, m(1, \eta), \epsilon) - f(0, m(0, \eta), \epsilon)\right) = E(y|d = 1) - E(y|d = 0).$$

This is the "overall" effect of the treatment on the outcome. This allows the individuals to choose other actions or choose a different network in the intervening stages. Now, consider the direct treatment effect, $DTE(G_0)$. The data reveal:

$$\begin{aligned} E(y|d = 1, G_0) = &\quad E(f(1, G_0, \epsilon)|m(1, \eta) = G_0) \\ E(y|d = 0, G_0) = &\quad E(f(0, G_0, \epsilon)|m(0, \eta) = G_0). \end{aligned}$$

It is difficult to learn anything beyond this, and especially it is difficult to learn the direct treatment effect, without further assumptions even when d is randomly assigned unless η is independent of ϵ. It possible to get the worst case bound as follows. For simplicity, suppose that $G \in \{G_0, G_1\}$. Then, using total probability, $E(f(1, G_0, \epsilon)) = \underline{E(f(1, G_0, \epsilon)|d = 1, G = G_0)P(1, G_0)} + E(f(1, G_0, \epsilon)|d = 1, G = G_1)\underline{P(1, G_1)} + E(f(1, G_0, \epsilon)|d = 0, G = G_0)\underline{P(0, G_0)} + E(f(1, G_0, \epsilon)|d = 0, G = G_1)\underline{P(0, G_1)}$, with a similar expression for $E(f(0, G_0, \epsilon))$, and only the underlined quantities can be point identified directly from the data. A similar analysis is also possible for $DTE(G_1)$.

There are other approaches to identification, motivated by the large literature on partial identification in treatment response models. See also Manski (2013) or Lazzati (2015). Simple bounds can be obtained via exclusion restrictions. Again, for simplicity, suppose that

$G \in \{G_0, G_1\}$ and suppose there exists a binary instrumental variable $Z \in \{Z_0, Z_1\}$ such that

$$G_i = m(d_i, Z_i, \eta_i).$$

Let the following exclusion restriction hold

$$f(d, G, Z, \epsilon) \equiv f(d, G, \epsilon) \quad \text{and} \quad (\epsilon, \eta) \perp (d, Z)$$

and assume the relevance condition that $m(d, Z, \eta)$ is a nontrivial function of Z. Then, the exclusion restriction and random assignment of treatment implies for any given s that

$$\begin{aligned} P(y(1, G_1) \leq s) &= P(y(1, G_1) \leq s | d = 1, Z) \\ &\leq P(y(1, G_1) \leq s | d = 1, G_1, Z) P(G_1 | d = 1, Z) + (1 - P(G_1 | d = 1, Z)) \\ &= P(y \leq s | d = 1, G_1, Z) P(G_1 | d = 1, Z) + (1 - P(G_1 | d = 1, Z)), \end{aligned}$$

which implies that

$$P(y(1, G_1) \leq s) \leq \min_{Z \in \{Z_0, Z_1\}} P(y \leq s | d = 1, G_1, Z) P(G_1 | d = 1, Z) + (1 - P(G_1 | d = 1, Z)).$$

And similarly,

$$P(y(1, G_1) \leq s) \geq \max_{Z \in \{Z_0, Z_1\}} P(y \leq s | d = 1, G_1, Z) P(G_1 | d = 1, Z).$$

Similar bounds can be derived on the distribution functions of $y(1, G_0)$, $y(0, G_0)$, and $y(0, G_1)$.

If point identification is required, it is possible to "reverse engineer" a parameter that is point identified. This is done through a choice model via the reduced forms:

$$\begin{aligned} E(y | d = 1, Z_1) &= E(f(1, G_1, \epsilon) 1 [\eta \in A_1(1, Z_1)] + f(1, G_0, \epsilon) 1 [\eta \in A_1^c(1, Z_1)]) \\ E(y | d = 1, Z_0) &= E(f(1, G_1, \epsilon) 1 [\eta \in A_1(1, Z_0)] + f(1, G_0, \epsilon) 1 [\eta \in A_1^c(1, Z_0)]) \end{aligned}$$

where $A_1(1, Z_1)$ and $A_1(1, Z_0)$ are the regions for η derived from the network formation model that result in the network G_1, at $Z = Z_1$ and $Z = Z_0$ respectively, when the treatment is $d = 1$. Similarly, $A_1(0, Z_1)$ and $A_1(0, Z_0)$ would be regions for when the treatment is $d = 0$. This can be equivalently written as:

$$\begin{aligned} E(y|d=1, Z_1) &= E((f(1, G_1, \epsilon) - f(1, G_0, \epsilon))1[\eta \in A_1(1, Z_1)]) + E(f(1, G_0, \epsilon)) \\ E(y|d=1, Z_0) &= E((f(1, G_1, \epsilon) - f(1, G_0, \epsilon))1[\eta \in A_1(1, Z_0)]) + E(f(1, G_0, \epsilon)), \end{aligned}$$

which yields

$$E(y|d=1, Z_1) - E(y|d=1, Z_0) = E\left((f(1, G_1, \epsilon) - f(1, G_0, \epsilon))\left(1[\eta \in A_1(1, Z_1)] - 1[\eta \in A_1(1, Z_0)]\right)\right).$$

Going further requires assumptions on the network formation model. One possibility is the model of Vytlacil (2002) (see also Imbens and Angrist (1994)) with a unidimensional unobservable (i.e., with monotonicity) as the network formation model. In particular, assume that $G_i = 1[\eta_i \leq g(d_i, Z_i)]$ where η_i is a scalar unobservable. Then, $A_1(d, Z_1) = \{\eta : \eta \leq g(d, Z_1)\}$ and $A_1(d, Z_0) = \{\eta : \eta \leq g(d, Z_0)\}$ and so taking for example the case of $g(1, Z_0) \leq g(1, Z_1)$,

$$1[\eta \in A_1(1, Z_1)] - 1[\eta \in A_1(1, Z_0)] = 1[g(1, Z_0) < \eta \leq g(1, Z_1)].$$

These can be known as the *complier networks* when treatment is $d = 1$. Hence,

$$E(f(1, G_1, \epsilon) - f(1, G_0, \epsilon)|g(1, Z_0) < \eta \leq g(1, Z_1)) = \frac{E(y|d=1, Z_1) - E(y|d=1, Z_0)}{P(G=G_1|d=1, Z_1) - P(G=G_1|d=1, Z_0)}$$

And so, this *local average indirect treatment effect* for *network compliers*, for when treatment is $d = 1$, is point identified. Similar steps can be taken for when treatment is $d = 0$.

It is not clear in general whether such a network formation model with a scalar unobservable is relevant. In particular, in networks with many decision makers, it is harder to obtain network formation models with scalar unobservables.

2.4.2. *Identification without random assignment of treatment.* Without random assignment of treatment, it is even more difficult to identify the direct and indirect treatment effects without further assumptions, because unobservables or confounders can be influencing both treatment selection and interim network formation or network changes. Aside from worst case bounds, one approach is to make assumptions from the empirical games literature. In particular, assume that there exists a binary variable $Z \in \{Z_0, Z_1\}$ such that

$$(Z = Z_1) \implies (G = G_1) \qquad \text{and} \qquad (Z = Z_0) \implies (G = G_0)$$

and Z is independent of (ϵ, η). This approximates the idea of "identification at infinity," whereby a certain explanatory variable can drive the decision making. For example, in models of airline markets, Z could be an indicator of whether a market is particularly profitable, and when that is the case a particular network arises (e.g., all potential entrants serve the market). These conditions render network formation exogenous conditional on these values of the variable Z, in which case the identification problem reverts back to a standard treatment response model. For example, without any further assumptions, it is possible to derive the standard worst case bounds on $DTE(G_0)$ and $DTE(G_1)$. These bounds can be tightened, for example by adding monotonicity assumptions on treatment response, like assuming that $y_i(1, G_0) \geq y_i(0, G_0)$ and $y_i(1, G_1) \geq y_i(0, G_1)$. It is also possible to have another instrument $W \in \{0, 1\}$, satisfying the exclusion restriction that

$$f(d, G, W, \epsilon) \equiv f(d, G, \epsilon) \quad \text{and} \quad ((\epsilon, \eta) \perp W)|Z,$$

in which case it is possible to identify the *local average direct treatment effect* for compliers in a model with a scalar/monotone unobservable. This use of double instruments (one for network formation and another for the usual treatment selection problem) can be used to identify the direct effect and the indirect effect.

2.5. **Other approaches to identification.** In some applications, detecting the existence (and direction) of a social interaction effect, but not the magnitude of the social interaction effect, may be an object of interest. It is possible to do so under weaker conditions than are used to identify the magnitude of the social interaction effect. In the context of various game theory models that could be used to model a social interaction, this identification problem has been addressed by de Paula and Tang (2012) and Kline (2016).

de Paula and Tang (2012) applies to incomplete information games. Consider the basic intuition for the de Paula and Tang (2012) identification strategy as applied to social interactions in the decision to smoke cigarettes. In equilibrium, the individuals will tend to either make similar (positively correlated) smoking decisions if there is a positive effect of a peer smoking on the utility from smoking, or dissimilar (negatively correlated) smoking decisions if there is a negative effect of a peer smoking on the utility from smoking. Hence, the direction of the peer effect is identified by the sign of the correlation in smoking decisions across peers.

Kline (2016) applies to complete information games. Consider the basic intuition for the Kline (2016) identification strategy as applied to social interactions in the decision to smoke cigarettes. Increasing the level of an excluded instrument that affects the utility an individual gets from smoking cigarettes but not the utility function of the other individual will (despite the complications in a game, like multiple equilibria) tend to increase the probability that individual smokes cigarettes. Then, as a consequence, in equilibrium, the other individual will either increase or decrease the probability of smoking cigarettes, depending on whether there is a positive effect of a peer smoking on the utility from smoking or a negative effect of a peer smoking on the utility from smoking.

Most identification analysis of models of social interactions, particularly in the context of continuous outcomes like variants of the linear-in-means models, focuses on mean outcomes. However, higher order moments of the outcomes, particular variances, can also be informative about social interactions. Graham (2008) shows that comparing between-group variation in outcomes, between large and small groups, to the within-group variation in outcomes, between large and small groups, is informative about social interaction effects, under suitable random assignment assumptions. The intuition for this identification strategy is that social interactions will tend to "amplify" the differences in outcomes when comparing between groups rather than comparing within groups.

It is also possible to use random or quasi-random variation to identify models of social interactions. For example, Sacerdote (2001) uses random assignment of roommates and dormmates to investigate social interactions in college GPA and other educational outcomes. Or for example, Imberman, Kugler, and Sacerdote (2012) uses inflows of evacuees from Hurricane Katrina into receiving schools to study impacts on incumbent students.

3. Specification of models of social interactions

3.1. Empirical individual treatment response.

An important consideration in the specification of an econometric model of social interactions is whether or not the model (and empirical setting) satisfies the assumption of *empirical individual treatment response* (EITR):

Definition 1. Consider a model with response functions $y_{ig}(\mathbf{y}_{-ig}, \mathbf{d}_g)$, which is the response of individual i in group g, where $\mathbf{y}_{-ig}$ is a possibly counterfactual specification of the outcomes of the other individuals in group g and $\mathbf{d}_g$ is a possibly counterfactual specification of the treatments of all individuals in group g. The model satisfies empirical individual treatment response, or EITR, if the data satisfy $y_{ig} = y_{ig}(Y_{-ig}, D_g)$, where y_{ig} is the actual outcome of individual i in group g, Y_{-ig} are the actual outcomes of the other individuals in group g, and D_g are the actual treatments of all individuals in group g.

This assumption relates the model to the data. As above, the response function is the object of interest, so the EITR assumption relates the data to the object of interest. If there were no social interaction, then this assumption is standard and uncontroversial: the response function at treatment d for individual i is observed if i is treated with d. In contrast, the assumption of EITR in models with social interaction entails implicit behavioral assumptions that are not necessarily implied by economic theory. Note that by construction, the specification of the linear-in-means model discussed in Section 2.1 satisfies EITR. Therefore, any criticism of models that satisfy EITR apply in particular to that specification of the linear-in-means model. The terminology *individual treatment response* refers to the fact that, from the perspective of each individual, the individual is "treated" by the treatments of all individuals in the group *and* the outcomes of the other individuals in the group.

Before analyzing the EITR assumption, it is necessary and nontrivial to make the assumption that such a response function even exists as an autonomous relationship between outcomes and treatments (e.g., Wooldridge (2010, Section 9.1)). Hence, in this chapter, and in analogy to the standard interpretation of response functions without social interactions, it is supposed that the response function gives the utility maximizing "response" of individual i in group g to being "treated" with the treatments of all individuals in the group and the outcomes of the other individuals in the group. This sort of response function was the focus of Kline and Tamer (2012), as discussed in Section 2.3.

EITR can be too strong in models with social interactions because it imposes the assumption that the observed outcome of each individual i is the response to the observed outcome of individuals $-i$. Viewing the model of the social interaction as a game theory model, so that the outcomes $\mathbf{y}_g$ are determined as the outcome of a game theory model, this is essentially the definition of pure strategy Nash equilibrium play with complete information.

It is well known that pure strategy Nash equilibrium with complete information often is not a reasonable characterization of actual behavior in some settings (e.g., Camerer (2003)). Even with the maintained assumption of complete information, many alternatives to pure strategy

Nash equilibrium play are entertained both in the theoretical game theory literature and also the experimental literature. Among many other examples this includes mixed strategy Nash equilibrium, rationalizability (i.e., Bernheim (1984) and Pearce (1984)), quantal response (i.e., McKelvey and Palfrey (1995)), and models of "bounded reasoning" (i.e., Camerer, Ho, and Chong (2004) and Costa-Gomes and Crawford (2006)). Conceivably, individuals might also randomize over the decision rule / solution concept that they use, as explored in Kline (2017). This would also lead to similar issues.

For example, if mixed strategies are used, then it is possible even in Nash equilibrium play for the condition that $y_{ig} = y_{ig}(Y_{-ig}, D_g)$ to be false, since with mixed strategies it is possible that the realizations of the mixed strategies do not (considered as pure strategies) comprise mutual best responses. In many games that could be used to model social interactions, it is not reasonable to rule out mixed strategies, because many games do not have a Nash equilibrium in pure strategies, and in others it may be "reasonable" that when there are multiple equilibria the mixed strategy equilibrium is selected.

The assumption of complete information is also a strong assumption. The assumption of EITR rules out incomplete information basically for the same reason that it rules out mixed strategies: if there is incomplete information then it is possible even in a Bayesian Nash equilibrium in which each type plays a pure strategy for the condition that $y_{ig} = y_{ig}(Y_{-ig}, D_g)$ to be false, since with incomplete information each individual is effectively interacting with other individuals using mixed strategies induced by the distribution over types. Therefore, the choice of y_{ig} in the data may not actually be made with knowledge of the choices of Y_{-ig} of the other individuals. In contrast, in models without social interactions, the econometrician observes the response to the realized treatment, for essentially any plausible specification of how that treatment is selected.

3.2. **Empirical group treatment response.** Another possible specification of the response function expresses the outcomes of all individuals in the group as a function of the exogenous treatments of all individuals in the group. Thus, generically, this specification of the response function $y_{ig}(\cdot)$ relates the outcome of individual i in group g to the treatment d_{ig} of individual i in group g and the treatments d_{-ig} of other individuals in group g.

This is relatively more similar to a standard response function in models without social interactions, but with vector outcomes and treatments. The model of the social interaction presumably places additional structure on this response function.

As with the assumption that the response function even existed in the previous section, it is an assumption that the social interaction process has a unique outcome, as a function of the profile of exogenous treatments. This specification of the response function does not reveal anything about the mechanism of the social interaction, but nevertheless can be the object of interest when the question is the relationship between exogenous treatments and outcomes. This sort of response function was the focus of Manski (2013) and Lazzati (2015), as discussed in Section 2.3.

The assumption of *empirical group treatment response* (EGTR) is the assumption that $Y_g = y_g(D_g)$ in the data. As with EITR, in models with social interaction this assumption involves implicit behavioral assumptions that are not necessarily implied by economic theory.

Definition 2. Consider a model with a group response function $y_g(\mathbf{d}_g)$, which is the response of all individuals in group g, where $\mathbf{d}_g$ is a possibly counterfactual specification of the treatments of all individuals in group g. The model is said to satisfy empirical group treatment response, or EGTR, if the data satisfies $Y_g = y_g(D_g)$, where Y_g are the actual outcomes of all individuals in group g and D_g are the actual treatments of all individuals in group g.

The terminology *group treatment response* refers to the fact that, from the perspective of the entire group, the group of individuals is "treated" by the treatments of all individuals in the group.

EGTR implicitly entails the assumption of a unique equilibrium of the social interaction process. Otherwise, if there are multiple equilibrium outcomes, there are multiple potential outcomes for the group for certain profiles of treatments **d**. And, indeed, in many settings of social interaction there are multiple equilibria according to reasonable solution concepts, like Nash equilibrium. Similarly, EGTR entails the assumption of the use of pure strategies in the social interaction process, although not necessarily in Nash equilibrium. If there were mixed strategies, then again there could be multiple potential outcomes even for a fixed profile of treatments **d**. Therefore, in general $y_g(\mathbf{d}_g)$ is a correspondence rather than a function, so $Y_g \in y_g(D_g)$ but not $Y_g = y_g(D_g)$. Note that the "reduced form" of the linear-in-means model in Equation 3 shows that the linear-in-means model does satisfy EGTR.

3.3. **Recap.** Consequently, in models with social interactions, EITR and EGTR can imply significant assumptions on behavior that are not necessarily implied by economic theory. Many models, including the linear-in-means model, satisfy these assumptions, and therefore imply these signification assumptions on behavior.

The next section illustrates these considerations by a concrete example of a setting with social interactions. Note that neither of EITR and EGTR necessarily implies the other. EITR can hold while EGTR fails if there is an interaction with pure strategy Nash equilibrium play with complete information and multiple equilibria. EGTR can hold while EITR fails if there is a unique potential outcome according to some solution concept that describes the behavior of the group, but the outcome is not comprised of mutual best responses considered as pure strategies. For example, if there is non-equilibrium behavior, then there can still be a unique potential outcome for the group so EGTR holds. But in general non-equilibrium behavior does not comprise mutual best responses considered as pure strategies, so EITR fails. EITR and EGTR can both hold if there is a unique pure strategy Nash equilibrium with complete information. And both can fail if there are multiple potential outcomes that do not comprise mutual best responses considered as pure strategies, which may happen when the equilibrium is in mixed strategies.

However, in addition to the considerations already discussed, there are settings in which EITR and/or EGTR may be satisfied. In particular, if there is perfect information, so that some individuals observe the decisions of the other individuals before making their own decisions, then it may be reasonable to assume EITR for some individuals. Treatment response models without social interaction can be viewed as a game with perfect information in which "nature" selects the treatments, and then the individuals "respond" to that treatment. This justifies the implicit assumption of EITR in models without social interactions.

3.4. **Immunization and infectious disease with social interaction.** Infections disease involves social interactions, and can be used to illustrate issues relating to the specification of models of social interactions. This section discusses two related models of infectious disease. The first is a model of the *decision to get immunized*, in which the object of interest is the response function that gives an individual's decision to immunize as a function of the immunization decisions by others in the individual's group. The second model is a model of *health outcomes under a policy intervention*, in which the object of interest is the health production function that relates an individual's health outcome to the treatments of everyone in that individual's group. Although the discussion relates to this specific application, the general issues raised would equally apply to other empirical settings.

3.4.1. *Models of the decision whether to get immunized.* Suppose the decision to get immunized made by individual i in group g, y_{ig}, is subject to a response function $y_{ig}(\mathbf{y}_{-ig})$, which relates the immunization decision of individual i in group g to the immunization decisions of the other individuals in group g. The econometrics problem is to learn about functionals of $y_{ig}(.)$ using a data set $(Y_g : g = 1, \ldots, G)$ of group immunization decisions. The response function is written as a function only of the immunization decision of others, but, as with all such models, can be made implicitly to depend on exogenous observed and unobserved variables (i.e., $y_{ig}(\mathbf{y}_{-ig}) \equiv y(\mathbf{y}_{-ig}, X_{ig}, \epsilon_{ig})$) through the indexing by i and g.

Consider the immunization decisions of two people in close contact who take as given the immunization decisions of the rest of their reference group. Similar ideas apply to immunization decisions of larger groups, like immunization decisions of students in a classroom or immunization decisions of residents of a local community. Since the two people are in close contact it is very likely the immunization status and health outcome of one individual affects the health outcomes of the other individual.

There is a private cost to getting immunized (e.g., the monetary cost, the time cost, or the health cost of perceived negative effects of immunization), but a group benefit. Therefore, the individuals may wish either to anti-coordinate or coordinate their immunization decisions, as described in the following two specifications. Such considerations would be important when considering an empirical analysis of immunization decisions, as they would impact the specification of the econometric model.

First, suppose that if at least one of the pair gets immunized they both "avoid" the disease and get β utils from health. However, there is a private cost of θ utils to getting the immunization. If neither gets immunized they both "get" the disease and get 0 utils from health. Assume that $0 < \theta < \beta$, and assume that these payoffs are common knowledge.

This interaction can be modeled by the normal form game specified in game 1a, where action 0 is not get immunized and action 1 is get immunized. This game has three Nash equilbria, which are reasonable candidates for predictions of the immunization decisions of the two individuals. There are two pure strategy equilibria in which one person gets immunized and the other doesn't, and there is a mixed strategy equilibrium in which both people play a strategy to get immunized with probability $\frac{\beta-\theta}{\beta}$. Hence, with these payoffs, there is a preference for anti-coordination. This basically results in the payoffs of the game of chicken, as described in Fudenberg and Tirole (1991, p. 18) among many places, where the "weak" action corresponds to getting immunized and the "tough" action corresponds to not getting immunized.

Second, suppose instead that both need to get immunized to avoid the illness. The payoffs are similar to the case above, but modified to account for the fact that now both individuals need to get immunized for there to be a health benefit. This is described in game 1b. This is basically a coordination game where both individuals get the same payoff to coordinating, and the payoff depends on what they coordinate on, after a payoff normalization. As with the previous specification, this game has three Nash equlbria. There are two pure strategy equilibria, one in which both individuals get immunized and one in which neither individual gets immunized, and there is a mixed strategy equilibrium in which both individuals play a strategy to get immunized with probability $\frac{\theta}{\beta}$.

	$y_2 = 0$	$y_2 = 1$
$y_1 = 0$	$0, 0$	$\beta, \beta - \theta$
$y_1 = 1$	$\beta - \theta, \beta$	$\beta - \theta, \beta - \theta$

(A) Anti-coordination

	$y_2 = 0$	$y_2 = 1$
$y_1 = 0$	$0, 0$	$0, -\theta$
$y_1 = 1$	$-\theta, 0$	$\beta - \theta, \beta - \theta$

(B) Coordination

Table 1. Two specifications of games for immunization decisions: (a) preference for anti-coordination and (b) preference for coordination

Similar models that suggest that immunization decisions involve strategic interaction, and in particular randomization of individual immunization decisions, include Bauch, Galvani, and Earn (2003), Bauch and Earn (2004), Galvani, Reluga, and Chapman (2007), Vardavas, Breban, and Blower (2007), and Bauch, Bhattacharyya, Ball, and Boni (2010). In the games described above the randomization is due to the possible use of mixed strategies, but there can also be randomization that is due to incomplete information, for example when individuals do not know the others' preferences over health and other goods. In that case, there is effectively randomization, from the perspective of each individual, induced by the distribution over types.

In addition to randomization of individual immunization decisions, it is also plausible in this sort of setting that there is not Nash equilibrium play. For example, in the case of preference for anti-coordination it is plausible that the outcome is that neither individual gets immunized, even though this is not a pure strategy Nash equilibrium. This is plausible because it could be that each individual believes the other individual will get immunized, which is a reasonable belief since getting immunized is the best response if the other individual does not get immunized, and so on. In other words, neither individual getting immunized is a rationalizable outcome. Consequently, neither individual getting immunized is a reasonable candidate for the prediction of the immunization decisions of the two individuals, even though it is not a pure strategy Nash equilibrium. This outcome could also arise from the realization of the mixed strategy Nash equilibrium.

Consider the econometrics problem of recovering information about this model of immunization decisions, based on data that consists of the immunization decisions of pairs of individuals. The assumption of EITR implies that, for example, for a pair of individuals in the data in which both are observed to get immunized, the response of each individual to the other individual getting immunized is to get immunized itself.

This rules out the possibility that the data comes from the mixed strategy equilibrium, or a rationalizable outcome. In either such case, and also for other reasons like the existence of incomplete information that also results in randomization, it need not be that immunization decisions Y_g in the data satisfy $y_{ig} = y_{ig}(Y_{-ig})$. Consequently, in those cases, EITR does not hold, and so using a model based on EITR can lead to misleading conclusions. Consider for example a public health organization that observes data on immunization decisions of married couples, and finds that fraction p of husbands whose wives are immunized also gets immunized himself, for some $p \in (0, 1)$. This is consistent with the mixed strategy Nash equilibrium behavior both when there is preference for anti-coordination and preference for coordination. The public health organization might conclude from this, based on an implicit or explicit assumption of EITR, that it is worthwhile to promote immunization among women, with the "understanding" from the data that the husbands would be reasonably likely to also get immunized if their wives get immunized, assuming that p is reasonably large. However, the validity of this conclusion depends on whether there is preference for anti-coordination or preference for coordination. In the first case this policy intervention would have the result of reducing the rate of immunization among men to 0 (among married couples in which the wife does get immunized) while in the second case it would have the result of increasing the rate of immunization among men to 1 (among married couples in which the wife does get immunized).

Obviously, these two scenarios have very different implications for the public health organization, but these considerations are essentially ignored when assuming EITR. In the case in which the immunization rate among men goes to 0, for example, it may be the case that this intervention is not worth the cost. The same reasoning holds for other policy interventions. For example, a public health organization might consider whether to promote vaccination among some subpopulation of a school (e.g., a particular grade level, or the teachers and staff), using the observed data to help predict the behavioral response of the rest of the school. The same considerations imply that assuming EITR can lead to misleading conclusions that the health organization draws from the data about such a policy intervention. It could be, for example, that the data would suggest that the result of this policy intervention is that, again, roughly fraction p of the rest of the school would get immunized, if a focal individual like an instructor is immunized, because that is observed in the data. But as before that analysis relies on behavioral assumptions that may not apply in all empirical settings.

3.4.2. *Models of health outcomes.* Now suppose that the object of interest is the "production function" of health outcomes as a function of the vector of treatments in a group. The treatments could be provision or subsidization of immunization, supply of bed nets, some sort of public health awareness campaign, or anything else that a social planner can manipulate that affects health outcomes. Let the health outcome of individual i in group g, h_{ig}, be given by $h_{ig} = h_{ig}(\mathbf{d}_g, \mathbf{A}_g)$ where $h_{ig}(\cdot)$ is a function, and $\mathbf{A}_g$ is an unobservable that represents actions that individuals take, possibly in response to the treatments $\mathbf{d}_g$. For example, given the treatments in $\mathbf{d}_g$, $\mathbf{A}_g$ can capture the decision of individuals in the group of whether (and how much) to interact with other individuals who may have the disease. If $\mathbf{A}_g$ is a *deterministic function* of $\mathbf{d}_g$ then it is without loss of generality to write that $h_{ig} = \upsilon_{ig}(\mathbf{d}_g) = h_{ig}(\mathbf{d}_g, \mathbf{A}_g(\mathbf{d}_g))$. Consequently, EGTR is plausible under this condition, because then the group health outcome is a unique function of the treatment, even though there is an intermediate behavior stage that also affects health outcomes. For recent work on identification in these settings of treatment response where an outcome is a function of the treatments of others, see Manski (2013) and Lazzati (2015), as discussed in Section 2.3.

There are two important cases where $\mathbf{A}_g$ is not a unique function of $\mathbf{d}_g$. The first case is when the behavioral stage of the model, considered as conditional on the "parameters" $\mathbf{d}_g$, can have multiple equilibria. In this case, $\mathbf{A}_g$ can take many different values for some $\mathbf{d}_g$. As a consequence, it is not credible to assume EGTR, because for a fixed vector of treatments, the group can have multiple different potential outcomes. The second case is when the behavioral stage of the model involves mixed strategies. Then again, $\mathbf{A}_g$ can take many different values for some $\mathbf{d}_g$.

The possibility of models of behavior with mixed strategies has already been discussed above. For example, suppose that $\mathbf{d}_g$ amounts to some sort of intervention that affects θ, the private util cost of immunization, in the model of Section 3.4.1. Then $\mathbf{A}_g$ is not a function. For example, $\mathbf{d}_g$ could be a subsidy to get vaccinated, and $\mathbf{A}_g$ could be some action that depends on $\mathbf{d}_g$ that affects payoffs, like whether to actually get immunized. In general, the health outcome differs depending on which actions are actually realized from the mixed strategies. For example, in both games, if it happens that both individuals get immunized, then neither individual gets the disease. But, if it happens that neither individual gets immunized, then both individuals get the disease. Consequently, under this model, there are multiple potential outcomes for the group given a fixed treatment $\mathbf{d}_g$. There are also multiple potential outcomes because of the existence of multiple equilibria in those games. For example, in the game with a preference for coordination, if it happens that the equilibrium in which both individuals get immunized is selected, then neither individual gets the disease. But, if it happens that the equilibrium in which neither individual gets immunized is selected, then both individuals get the disease.

Further, for another example of a social interaction with multiple equilibria, consider models of disease in which individuals can modify their behavior (denoted above as $\mathbf{A}_g$) as a function of the prevalence of the disease. For example, in the case of a sexually transmitted infection (i.e., Kremer (1996)) this can be the number of sexual partners per period. More generally, the behavior can be the decisions relating to interaction with other individuals who may have the disease. As before, $\mathbf{d}_g$ is a vector of treatments like those discussed above, like subsidization of some sort of medical care. The actions $\mathbf{A}_g$ are the decisions of individuals about interacting with others. Kremer (1996) shows that these models can have multiple equilibria that have different prevalences of the disease. Consequently, the assumption of EGTR rules out this class of model of disease and behavior. The intuition for existence of multiple equilbria in such settings is described in detail in Kremer (1996).

Even though the vector of treatments is exogenous to the model of health outcomes, the assumption of EGTR rules out models which have the potential for multiple equilibria. The assumption of EGTR also rules out mixed strategies. All of these considerations apply equally to observational data and experimental data. There are many advantages of experiments, for example because experiments can solve endogeneity problems. Experimental data and econometric modeling of social interaction provide complementary features to an analysis. The experiment helps to solve issues with endogeneity with respect to the relation between observable and unobservable characteristics of the individuals, while the econometric modeling helps resolve the issues discussed in this section, that are unrelated to the endogeneity problem, but are related to how to link data and models in settings where social interactions are important.

4. Policy relevance of models of social interactions

Models of social interactions are often used to make policy relevant claims like that peers "cause each other" to have certain outcomes, as in claims like that "a student who has high achieving peers has increased educational achievement itself." In some settings it is plausible that a social planner could directly manipulate the outcome of some individuals, for example by preventing some actions for some individuals as in drug and alcohol policy, or by a mandatory immunization policy. However, in general, because the outcomes in models of social interactions are simultaneously determined, it may not be possible for a social planner to directly manipulate the outcomes of individuals. This suggests that caution is warranted when interpreting the results of estimating models of social interactions, particularly when making claims about the effect of the outcomes on other outcomes. It may be more straightforward to investigate the effect of manipulating the exogenous explanatory variables, accounting for the effects of the social interaction. For example, in the context of the linear-in-means model, the effect of the exogenous explanatory variables is given by Equation 4. Similar ideas were addressed in Kline and Tamer (2013). Beyond the specific context of the linear-in-means model, these considerations suggest the role of the response functions discussed in Section 2.3, that are the "reduced form" that map the treatments of all individuals to the outcomes of all individuals. Note that it may not be desirable to simply run a linear regression of "outcomes" on "exogenous explanatory variables" even if the reduced form is the object of interest, because the reduced form of a social interactions model is not necessarily a linear model, as demonstrated for the linear-in-means model in Equation 3. In particular, even for the linear-in-means model with a linear structural form, the reduced form depends in a complicated non-linear way on the network G and parameter δ.

Nevertheless, the "interaction effects" can be useful to understand the "partial" effect of outcomes on other outcomes, even if outcomes cannot actually be directly manipulated. The response functions can be useful to get a sense of the social interaction mechanism when considering a policy intervention that directly affects the outcomes of individuals privately, for example some treatment (e.g., subsidized immunization), but because of the social interaction process may indirectly affect the outcomes of everyone in the group.

It is also useful to consider possible "re-equilibrization" effects of policy interventions or other manipulations of outcomes. In particular, if a particular outcome is somehow directly manipulated, then the outcomes of other individuals in the group will change in response, according to the equilibrium of the social interaction. Then, unless the originally manipulated outcome is still held fixed by the manipulation, the original outcome might change in equilibrium in response to the changes of the outcomes of the other individuals. In other words, there can be a difference between a one-time manipulation of an outcome, in which case that outcome might subsequently change due to re-equilibrization, and a permanent manipulation of an outcome to a fixed outcome, in which case that outcome is "removed" from the model of the social interaction and no longer is "re-equilibriated."

Particularly in the context of social interactions models with an underlying network structure of interactions within groups, an important consideration is the joint determination of outcomes and the network. Interventions, whether they be related to manipulating outcomes directly or the exogenous explanatory variables, can be suspected to also have effects on the network. If so, then counterfactual predictions of the outcomes after the intervention may be poor predictions if they assume that the network will be unaffected by the intervention, although such a counterfactual that "holds fixed" the network can still be useful to understand the "partial" effect of the intervention on outcomes, even if it is not predictive of the actual outcomes. This is the direct effect of the intervention. For example, it could be possible to use a model of network formation to predict the resulting network after the intervention, and then use that predicted model in the model of social interactions to predict the resulting outcomes after the intervention. Carrell, Sacerdote, and West (2013) find evidence that such considerations can be practically relevant, in that when they attempted to manipulate peer groups, individuals formed a network of connections within each manipulated peer group that lead to individuals avoiding interacting with individuals that the manipulated peer group had been designed for them to interact with.

It is also useful to understand the mediating effect of networks in models of social interactions. Does a particular policy intervention have different effects in different groups (i.e., cities/villages/neighborhoods) that have different networks of connections? In many applications, the network will be sufficiently "stable" that a relatively minor intervention affecting one outcome (out of many outcomes affected by the network) may be reasonably assumed to have negligible effect on the network. Moreover, it is useful to measure how much of the change in outcomes is due to changes in the network. This is the indirect effect of the intervention. It may be possible to parametrize the link between the treatment and the network and outcomes and use a full model to learn all the causal effects of interest.

References

Aradillas-Lopez, A. (2010): "Semiparametric estimation of a simultaneous game with incomplete information," *Journal of Econometrics*, 157(2), 409–431.

Aradillas-Lopez, A. (2012): "Pairwise-difference estimation of incomplete information games," *Journal of Econometrics*, 168(1), 120–140.

Aradillas-Lopez, A., and E. Tamer (2008): "The identification power of equilibrium in simple games," *Journal of Business & Economic Statistics*, 26(3), 261–283.

Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2010): "Estimating static models of strategic interactions," *Journal of Business & Economic Statistics*, 28(4), 469–482.

Bajari, P., H. Hong, and S. P. Ryan (2010): "Identification and estimation of a discrete game of complete information," *Econometrica*, 78(5), 1529–1568.

Bauch, C., S. Bhattacharyya, R. Ball, and M. Boni (2010): "Rapid emergence of free-riding behavior in new pediatric immunization programs," *PLoS ONE*, 5(9), 505–514.

Bauch, C., and D. Earn (2004): "Vaccination and the theory of games," *Proceedings of the National Academy of Sciences of the United States of America*, 101(36), 13391–13394.

Bauch, C., A. Galvani, and D. Earn (2003): "Group interest versus self-interest in smallpox vaccination policy," *Proceedings of the National Academy of Sciences of the United States of America*, 100(18), 10564–10567.

Beresteanu, A., I. Molchanov, and F. Molinari (2011): "Sharp identification regions in models with convex moment predictions," *Econometrica*, 79(6), 1785–1821.

Bernheim, B. (1984): "Rationalizable strategic behavior," *Econometrica*, 52(4), 1007–1028.

Berry, S., and P. Reiss (2007): "Empirical models of entry and market structure," in *Handbook of Industrial Organization*, ed. by M. Armstrong, and R. Porter, vol. 3, pp. 1845–1886. North-Holland, Amsterdam.

Berry, S., and E. Tamer (2006): "Identification in models of oligopoly entry," in *Advances in Economics and Econometrics Theory and Applications, Ninth World Congress*, ed. by R. Blundell, W. K. Newey, and T. Persson, pp. 46–85. Cambridge University Press, Cambridge.

Berry, S. T. (1992): "Estimation of a model of entry in the airline industry," *Econometrica*, 60(4), 889–917.

Blume, L. E., W. A. Brock, S. N. Durlauf, and Y. M. Ioannides (2011): "Identification of Social Interactions," in *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. O. Jackson, vol. 1, pp. 853 – 964. North-Holland, Amsterdam.

Blume, L. E., W. A. Brock, S. N. Durlauf, and R. Jayaraman (2015): "Linear social interactions models," *Journal of Political Economy*, 123(2), 444–496.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009): "Identification of peer effects through social networks," *Journal of Econometrics*, 150(1), 41–55.

Bresnahan, T., and P. Reiss (1991): "Empirical models of discrete games," *Journal of Econometrics*, 48(1-2), 57–81.

Brock, W. A., and S. N. Durlauf (2001): "Discrete choice with social interactions," *The Review of Economic Studies*, 68(2), 235–260.

Brock, W. A., and S. N. Durlauf (2007): "Identification of binary choice models with social interactions," *Journal of Econometrics*, 140(1), 52–75.

Camerer, C. (2003): *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, Princeton, NJ.

Camerer, C., T. Ho, and J. Chong (2004): "A cognitive hierarchy model of games," *Quarterly Journal of Economics*, 119(3), 861–898.

Card, D., and L. Giuliano (2013): "Peer effects and multiple equilibria in the risky behavior of friends," *Review of Economics and Statistics*, 95(4), 1130–1149.

Carrell, S. E., B. I. Sacerdote, and J. E. West (2013): "From natural variation to optimal policy? The importance of endogenous peer group formation," *Econometrica*, 81(3), 855–882.

Christakis, N. A., and J. H. Fowler (2007): "The spread of obesity in a large social network over 32 years," *New England Journal of Medicine*, 357(4), 370–379.

Ciliberto, F., and E. Tamer (2009): "Market structure and multiple equilibria in airline markets," *Econometrica*, 77(6), 1791–1828.

Costa-Gomes, M., and V. Crawford (2006): "Cognition and behavior in two-person guessing games: an experimental study," *American Economic Review*, 96(5), 1737–1768.

Davezies, L., X. d'Haultfoeuille, and D. Fougère (2009): "Identification of peer effects using group size variation," *The Econometrics Journal*, 12(3), 397–413.

De Giorgi, G., M. Pellizzari, and S. Redaelli (2010): "Identification of social interactions through partially overlapping peer groups," *American Economic Journal: Applied Economics*, 2(2), 241–275.

Durlauf, S. N., and Y. M. Ioannides (2010): "Social interactions," *Annual Review of Economics*, 2, 451–478.

Epple, D., and R. Romano (2011): "Peer effects in education: A survey of the theory and evidence," in *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. O. Jackson, vol. 1, pp. 1053–1163. North-Holland, Amsterdam.

Fox, J. T., and N. Lazzati (2016): "A note on identification of discrete choice models for bundles and binary games."

Fudenberg, D., and J. Tirole (1991): *Game Theory*. MIT Press, Cambridge, MA.

Galvani, A., T. Reluga, and G. Chapman (2007): "Long-standing influenza vaccination policy is in accord with individual self-interest but not with the utilitarian optimum," *Proceedings of the National Academy of Sciences of the United States of America*, 104(13), 5692–5697.

Glaeser, E., and J. Scheinkman (2001): "Measuring social interactions," in *Social Dynamics*, ed. by S. N. Durlauf, and H. P. Young, pp. 83–132. MIT Press, Cambridge, MA.

Graham, B. S. (2008): "Identifying social interactions through conditional variance restrictions," *Econometrica*, 76(3), 643–660.

Graham, B. S. (2015): "Methods of identification in social networks," *Annual Review of Economics*, 7, 465–485.

Graham, B. S., and J. Hahn (2005): "Identification and estimation of the linear-in-means model of social interactions," *Economics Letters*, 88(1), 1–6.

Grieco, P. L. (2014): "Discrete games with flexible information structures: An application to local grocery markets," *The RAND Journal of Economics*, 45(2), 303–340.

Honoré, B. E., and Á. de Paula (2010): "Interdependent durations," *The Review of Economic Studies*, 77(3), 1138–1163.

Honoré, B. E., and Á. de Paula (2011): "Interdependent durations in joint retirement."

Imbens, G. W., and J. D. Angrist (1994): "Identification and estimation of local average treatment effects," *Econometrica*, 62(2), 467–475.

Imberman, S. A., A. D. Kugler, and B. I. Sacerdote (2012): "Katrina's children: Evidence on the structure of peer effects from hurricane evacuees," *The American Economic Review*, 102(5), 2048–2082.

Kline, B. (2015): "Identification of complete information games," *Journal of Econometrics*, 189(1), 117–131.

Kline, B. (2016): "The empirical content of games with bounded regressors," *Quantitative Economics*, 7(1), 37–81.

Kline, B. (2017): "An empirical model of non-equilibrium behavior in games."

Kline, B., and E. Tamer (2012): "Bounds for best response functions in binary games," *Journal of Econometrics*, 166(1), 92–105.

Kline, B., and E. Tamer (2013): "Comment," *Journal of Business & Economic Statistics*, 31(3), 276–279.

Kline, B., and E. Tamer (2016): "Bayesian inference in a class of partially identified models," *Quantitative Economics*, 7(2), 329–366.

Krauth, B. V. (2006): "Simulation-based estimation of peer effects," *Journal of Econometrics*, 133(1), 243–271.

Kremer, M. (1996): "Integrating behavioral choice into epidemiological models of AIDS," *The Quarterly Journal of Economics*, 111(2), 549–573.

Kremer, M., and D. Levy (2008): "Peer effects and alcohol use among college students," *The Journal of Economic Perspectives*, 22(3), 189–189.

Lazzati, N. (2015): "Treatment response with social interactions: Partial identification via monotone comparative statics," *Quantitative Economics*, 6(1), 49–83.

Lee, L.-f. (2007): "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects," *Journal of Econometrics*, 140(2), 333–374.

Lee, L.-f., J. Li, and X. Lin (2014): "Binary choice models with social network under heterogeneous rational expectations," *Review of Economics and Statistics*, 96(3), 402–417.

Lewbel, A., and X. Tang (2015): "Identification and estimation of games with incomplete information using excluded regressors," *Journal of Econometrics*, 189(1), 229–244.

Liu, X., E. Patacchini, and Y. Zenou (2014): "Endogenous peer effects: local aggregate or local average?," *Journal of Economic Behavior & Organization*, 103, 39–59.

Lundborg, P. (2006): "Having the wrong friends? Peer effects in adolescent substance use," *Journal of Health Economics*, 25(2), 214–233.

Manski, C. (1993): "Identification of endogenous social effects: the reflection problem," *The Review of Economic Studies*, 60(3), 531–542.

Manski, C. F. (1997): "Monotone treatment response," *Econometrica*, 65(6), 1311–1334.

Manski, C. F. (2000): "Economic analysis of social interactions," *The Journal of Economic Perspectives*, 14(3), 115–136.

Manski, C. F. (2013): "Identification of treatment response with social interactions," *The Econometrics Journal*, 16(1), S1–S23.

Manski, C. F., and J. V. Pepper (2000): "Monotone instrumental variables: with an application to the returns to schooling," *Econometrica*, 68(4), 997–1010.

Manski, C. F., and J. V. Pepper (2009): "More on monotone instrumental variables," *The Econometrics Journal*, 12(s1), S200–S216.

Masten, M. (2017): "Random coefficients on endogenous variables in simultaneous equations models."

McKelvey, R., and T. Palfrey (1995): "Quantal response equilibria for normal form games," *Games and Economic Behavior*, 10(1), 6–38.

Moffitt, R. A. (2001): "Policy interventions, low-level equilibria, and social interactions," in *Social Dynamics*, ed. by S. N. Durlauf, and H. P. Young, chap. 3, pp. 45–82. MIT Press, Cambridge, MA.

Okumura, T., and E. Usui (2014): "Concave-monotone treatment response and monotone treatment selection: With an application to the returns to schooling," *Quantitative Economics*, 5(1), 175–194.

de Paula, A. (2009): "Inference in a synchronization game with social interactions," *Journal of Econometrics*, 148(1), 56–71.

de Paula, Á. (2016): "Econometrics of network models."

de Paula, Á., and X. Tang (2012): "Inference of signs of interaction effects in simultaneous games with incomplete information," *Econometrica*, 80(1), 143–172.

Pearce, D. (1984): "Rationalizable strategic behavior and the problem of perfection," *Econometrica*, 52(4), 1029–1050.

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

Sacerdote, B. (2001): "Peer effects with random assignment: Results for Dartmouth roommates," *The Quarterly Journal of Economics*, 116(2), 681–704.

Sacerdote, B. (2011): "Peer effects in education: How might they work, how big are they and how much do we know thus far?," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, S. J. Machin, and L. Woessmann, vol. 3, pp. 249–277. Elsevier, Amsterdam.

Scheinkman, J. A. (2008): "Social interactions," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf, and L. E. Blume, vol. 2. Palgrave Macmillan, London.

Soetevent, A. R., and P. Kooreman (2007): "A discrete-choice model with social interactions: with an application to high school teen behavior," *Journal of Applied Econometrics*, 22(3), 599–624.

Tamer, E. (2003): "Incomplete simultaneous discrete response model with multiple equilibria," *Review of Economic Studies*, 70(1), 147–165.

Tang, X. (2010): "Estimating simultaneous games with incomplete information under median restrictions," *Economics Letters*, 108(3), 273–276.

Vardavas, R., R. Breban, and S. Blower (2007): "Can influenza epidemics be prevented by voluntary vaccination?," *PLoS Computational Biology*, 3(5), e85.

Vytlacil, E. (2002): "Independence, monotonicity, and latent index models: An equivalence result," *Econometrica*, 70(1), 331–341.

Wan, Y., and H. Xu (2014): "Semiparametric identification of binary decision games of incomplete information with correlated private signals," *Journal of Econometrics*, 182(2), 235–246.

Wooldridge, J. M. (2010): *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, MA, 2 edn.

Xu, H. (2011): "Social interactions: A game theoretical approach."