Identification of Anonymous Endogenous Interactions

Charles F. Manski

SFI WORKING PAPER: 1996-04-019

SFI Working Papers contain accounts of scientific work of the author(s) and do not necessarily represent the views of the Santa Fe Institute. We accept papers intended for publication in peer-reviewed journals or proceedings volumes, but not papers that have already appeared in print. Except for papers by our external faculty, papers must be based on work done at SFI, inspired by an invited visit to or collaboration at SFI, or funded by an SFI grant.

©NOTICE: This working paper is included by permission of the contributing author(s) as a means to ensure timely distribution of the scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the author(s). It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author’s copyright. These works may be reposted only with the explicit permission of the copyright holder.

www.santafe.edu
IDENTIFICATION OF ANONYMOUS ENDOGENOUS INTERACTIONS

Charles F. Manski
Department of Economics
University of Wisconsin-Madison

September 1995

This research is supported by National Science Foundation Grant SBR92-23220.

1. Introduction

In theoretical studies of social interactions, we hypothesize a process and seek to deduce the implied outcomes. In inferential studies, we face an inverse logical problem. Given observations of outcomes and maintained assumptions, we seek to deduce the actual process generating the observations.

Econometricians have long found it useful to separate inferential problems into statistical and identification components (see Koopmans, 1949). Studies of identification seek to characterize the conclusions that could be drawn if one could use a given sampling process to obtain an unlimited number of observations. Studies of statistical inference seek to characterize the generally weaker conclusions that can be drawn from a finite number of observations. Analysis of identification logically comes first. Negative identification findings imply that statistical inference is fruitless: it makes no sense to try to use a sample of finite size to infer something that could not be learned even if a sample of infinite size were available. Positive identification findings imply that one should go on to examine the feasibility of statistical inference.

Throughout the modern development of the social sciences, analysis of the problem of empirical inference on social interaction processes has lagged far behind our willingness and ability to theorize about these processes. This asymmetry is
unfortunate. Theoretical studies hypothesizing alternative realities are ultimately sterile if we do not also ask how these alternative realities are empirically distinguishable.

The classic economic example of the asymmetric state of theory and inference is the analysis of competitive markets. The theory of equilibrium in markets with price-taking consumers and price-taking (or quantity-taking) firms was reasonably well-understood a century ago, but the corresponding econometric problem of identification of demand and supply from observations of equilibrium market transactions was only dimly understood until the 1940s (see Hood and Koopmans, 1953). Even then, the econometric literature developed under the strong maintained assumption that demand and supply are linear functions of price and quantity. The present-day econometric literature offers only scattered findings about the identifiability of nonlinear demand and supply functions. See Fisher (1966, Chapter 5), Roehrig (1988), and Manski (1995, Section 6.4).¹

Several years ago, I became aware of another sharp contrast between the state of theory and inference. The story begins with an empirical regularity: It is often observed that persons belonging to the same group tend to behave similarly.

For at least fifty years, sociologists and social psychologists have sought to explain this regularity by theorizing the existence of social interactions in which the behavior of an individual varies with the distribution of behavior in a "reference group" containing the individual. Depending on the context, these interactions may be called "social norms," "peer influences," "neighborhood effects," "conformity," "imitation," "contagion," "epidemics," "bandwagons," or "herd behavior." See, for example, Hyman (1942), Merton (1957), and Granovetter (1979).

¹ It is important to recognize that the work on estimation of nonlinear simultaneous equations that began in the 1970s and matured into the literature on method-of-moments estimation in the 1980s does not address the question of identification. This literature assumes identification and then focuses on the statistical aspects of inference (see Manski, 1988, pp. 93-94).
Economists have increasingly contributed to this theory, and have done much to formalize it in a coherent manner. See, for example, Schelling (1971), Pollak (1976), Jones (1984), Benabou (1992), Bernheim (1994), and Brock and Durlauf (1995). Economists have also become aware that various economic and physical processes share the same formal structure. Oligopoly models posit reaction functions, wherein the output of a given firm is a function of aggregate industry output. Endogenous growth models assume dynamic learning processes in which the output of a given firm depends on the lagged value of aggregate industry output. Biological models of contagion suppose that the probability an uninfected person becomes infected varies with the fraction of the population already infected. In physics, mean-field theory explains magnetism by hypothesizing that the orientation of a particle in space varies with the mean orientation of surrounding particles.

I shall use the term anonymous endogenous interactions to describe the class of social processes in which the behavior of an individual varies with the distribution of behavior in a group containing the individual. The interactions are endogenous because the outcome of each group member varies with the outcomes of the other group members, not with other attributes of the group. In contrast, the sociological literature on contextual interactions assumes that individual outcomes are influenced by the exogenous characteristics of the group members (see Sewell and Armer, 1966). The interactions are anonymous because they may be described without naming the members of the group or otherwise specifying the internal structure of the group. What matters is the distribution of outcomes in the group, not who experiences what outcome. In contrast, the interactions studied in the literatures on social networks and on local interactions require specification of the internal structure of the group.

Despite the large body of theory on anonymous endogenous interactions, I
found no work on the corresponding problem of identification from observations of outcomes. So I set out to study the problem (Manski, 1993). In particular, I wanted to understand the conditions under which these interactions can be distinguished empirically from other hypotheses that have often been advanced to explain the empirical regularity that individuals belonging to the same group tend to behave similarly. Three familiar hypotheses are

**endogenous interactions**, wherein the propensity of an individual to behave in some way varies with the behavior of the group.

**contextual** (exogenous) interactions, wherein the propensity of an individual to behave in some way varies with the exogenous characteristics of the group.

**correlated effects**, wherein individuals in the same group tend to behave similarly because they have similar individual characteristics or face similar institutional environments.

To illustrate, consider the determination of achievement in school. We have endogenous interactions if, all else equal, individual achievement tends to vary with the average achievement of the students in the youth's school, ethnic group, or other reference group. We have contextual interactions if achievement tends to vary with the socioeconomic composition of the reference group. We have correlated effects if youth in the same school tend to achieve similarly because they have similar family backgrounds or because they are taught by the same teachers.

In Manski (1993) I showed that, given the maintained assumption of a linear-in-means interaction model, a researcher who observes equilibrium outcomes and the composition of reference groups cannot empirically distinguish endogenous
interactions from contextual interactions or correlated effects. The researcher may or may not be able to determine whether there is some form of social interaction at work. Furthermore, a researcher who observes equilibrium outcomes but not the composition of reference groups cannot distinguish among a host of alternative interaction models, all of which hold tautologically.

I paraphrase these findings in Section 2. I then extend the analysis in some respects in Section 3. Section 4 gives conclusions. Here is the main notation used throughout the paper:

\[ t = 0, \ldots, T \]  
\[ x \]  
\[ w_t \ (z_t, u_t) \]  
\[ y_t \]  
\[ P(x, w_t, y_t, t = 0, \ldots, T) \]  
\[ P(y_{t \mid x}) \]  
\[ P(w_{t \mid x}) \]

The reader should take note of two maintained assumptions embedded in this notation. First, outcomes and covariates are potentially time-varying but reference groups are not. Second, each individual belongs to exactly one reference group, not to multiple reference groups. In other words, the population is assumed to partition into a set of mutually exclusive and exhaustive groups.

2. A Linear-in-Means Model

In this section, we assume that the outcomes at each date \( t = 1, \ldots, T \) are
generated by the *linear-in-means* model

\[(1a) \ y_t = a + \beta E(y_{t-1|x}) + E(z_{t-1|x})' + z_t' + u_t, \]

\[(1b) \ E(u_{t|x}, z_t) = x'\partial, \]

or, equivalently,

\[(2) \ E(y_{t|x}, z_t) = a + \beta E(y_{t-1|x}) + E(z_{t-1|x})' + x'\partial + z_t'. \]

Here \(y_t\) and \(u_t\) are scalars, while \(x\) and \(z_t\) are vectors. The scalar parameter \(\beta\) measures the strength of the endogenous interaction among the members of group \(x\), which is transmitted through the lagged group-mean outcome \(E(y_{t-1|x})\). The parameter vector \(\_\) measures the strength of the contextual interaction among the members of group \(x\), which is transmitted through the lagged group-mean covariate vector \(E(z_{t-1|x})\). The parameter vector \(\partial\) measures the direct variation of outcomes with \(x\), and so expresses the idea of correlated effects. The parameter vector \(\_\) measures the direct variation of outcomes with the contemporaneous individual covariates \(z_t\).

Integrating (2) with respect to \(P(z_{t|x})\), we find that the law of motion for \(E(y_{t|x})\) is

\[(3) \ E(y_{t|x}) = a + \beta E(y_{t-1|x}) + E(z_{t-1|x})' + x'\partial + E(z_{t|x})'. \]

If \(\beta < 1\) and the group-mean covariate vector \(E(z_{x})\) is time-invariant, there is a unique steady state, or *social equilibrium*, namely

\[(4) \ E(y_{x}) = a/(1 - \beta) + E(z_{x})(\_ + \_)/(1 - \beta) + x'\partial/(1 - \beta). \]
Now consider the inferential problem faced by a researcher who knows that outcomes satisfy (2), who observes a random sample of realizations of \((x, z_t, y_t, t = 0, \ldots, T)\) but not \(u_t\) and who wants to infer the parameters. Given minimal regularity conditions, the observational evidence enables the researcher to consistently estimate the conditional expectations \([E(y_{t|x}, z_t), E(y_t|x), E(z_{t|x})]\) at each value of \((x, z_t)\) realized in the population (technically, on the support of the regressors).\(^2\)

I want to focus on the identification component of the inferential problem and abstract from statistical concerns, so let us suppose that the researcher knows these conditional expectations. In particular, I want to focus on the situation in which \(\beta = 1\), \(E(z|x)\) is time-invariant, and each group \(x\) is observed in its unique equilibrium. Inspection of (4) shows that the equilibrium group-mean outcome \(E(y|x)\) is a linear function of the vector \([1, E(z|x), x]\); that is, these regressors are perfectly collinear. This implies that the parameters \((a, \beta, \_\_\_\, \delta)\) are unidentified. Thus, endogenous effects cannot be distinguished from contextual effects or from correlated effects.\(^3\)

One may or may not be able to learn if there is some form of social interaction. Inserting (4) into (2) yields the linear reduced form

\[
E(y_{-x}, z) = a/(1 - \beta) + E(z|x)[(\_\_\_\_\_\, \beta)/(1 - \beta)] + x'\delta/(1 - \beta) + z'
\]

\(^2\) In particular, consistent estimation is possible if these conditional expectations vary continuously with \((x, z_t)\), or if \((x, z_t)\) has discrete support.

\(^3\) This finding does not imply that the three hypotheses are indistinguishable in principle. They are distinguishable if groups are observed out-of-equilibrium or if an environmental or policy shift should change the parameters of model (2). To illustrate the implications of a parameter change, consider again the determination of achievement in school. Suppose that a technological advance in instruction increases the value of \(a\) to, say, \(a + k\), where \(k > 0\). By (4), equilibrium mean achievement rises by \(k/(1 - \beta)\) rather than by \(k\). Thus, endogenous interactions generate a "social multiplier" that is not generated by contextual interactions or correlated effects.
The solutions to (5) contain a unique value of the composite parameter \((\_ + \_\beta)/(1 - \beta)\) if \(E(z_x)\) is linearly independent of the vector \((1, x, z)\). In particular, \(E(z_x)\) must vary nonlinearly with \(x\). Although identification of \((\_ + \_\beta)/(1 - \beta)\) does not enable one to distinguish between endogenous and contextual interactions, it does enable one to learn if some social interaction is present. If the composite parameter is nonzero, then either \(\beta\) or \(\_\) must be nonzero.

The foregoing assumes that the researcher observes the reference groups into which the population is sorted. Suppose that these groups are unobserved. Can the researcher infer the composition of reference groups from observations of the outcomes \(y_t\) and covariates \(z_t\)?

The answer is necessarily negative if \(\beta = 1\), \(z\) is time-invariant (not just \(E(z_x)\)), and each group \(x\) is observed in its unique equilibrium. To see this, consider any hypothesized specification of reference groups such that \(z\) is a function of \(x\), say \(z = z(x)\). Then the equality

\[
E[y_x, z(x)] = E(y_x)
\]

holds tautologically. Comparing equations (6) and (2) shows that, with \(E(y_x)\) being time-invariant in equilibrium, (6) is the special case of (2) in which \(\beta = 1\) and \(a = \_ = \_\beta = \_\) = 0.

3. Pure Endogenous Interactions Models

Let \(Y\) denote the space of possible outcome values \(y\), \(W\) the space of possible covariate values \(w\), and \(G\) the space of all distributions on \(Y\). In this section, we assume that the outcomes at each date \(t = 1, \ldots, T\) are generated by a response function
of the form

\begin{equation}
    y_t = f[P(y_{t-1|x}), w_t],
\end{equation}

where \( f(\bullet, \bullet) : G \times W \rightarrow Y \). Thus, we assume that an individual's outcome \( y_t \) is determined in part by anonymous endogenous interactions, acting through the lagged group outcome distribution \( P(y_{t-1|x}) \), and in part by the individual's contemporaneous covariates \( w_t \). The members of group \( x \) may have similar covariates, so the model incorporates the idea of correlated effects. Contextual interactions, however, are assumed absent.

The time path of the group outcome distribution \( P(y_{t|x}) \) is determined by the form of \( f(\bullet, \bullet) \), by the initial condition \( P(y_{0|x}) \), and by the time path of the covariate distribution \( P(w_{t|x}) \). Let \( A \) denote a measurable subset of \( Y \). The law of motion for \( P(y_{t|x}) \) is

\begin{equation}
    P(y_t \in A | x) = P[w_t: f[P(y_{t-1|x}), w_t] \in A | x], \quad \text{all } A \subset Y.
\end{equation}

Given a time-invariant group covariate distribution \( P(w_x) \), a time-invariant outcome distribution \( P(y_x) \) is a social equilibrium if \( P(y_x) \) solves the system of equations

\begin{equation}
    P(y \in A | x) = P[w: f[P(y_x), w] \in A | x], \quad \text{all } A \subset Y.
\end{equation}

The set of equilibria is determined by the form of \( f(\bullet, \bullet) \) and by the covariate distribution \( P(w_x) \). The process may have a unique equilibrium in group \( x \), multiple equilibria, or no equilibrium at all.

3.1. Some Illustrations
Before discussing inference, I describe some models that illustrate alternative forms of the response function. These models express a range of ideas about the nature of interactions, yet all are quite simple.

*Linear-In Mean Model:* Let us begin with a modest variation on the model examined in Section 2. We assume that a scalar outcome varies linearly with the lagged group-mean outcome and separably with the individual's covariates. Thus

$$\text{(10)} \quad y_t = \beta E(y_{t-1|x}) + g(w_t).$$

As before, $\beta$ expresses the strength of the endogenous interaction. Now $g(\bullet): W \to \mathbb{R}$ expresses the manner in which covariates influence outcomes.

The law of motion for $E(y_{t|x})$ is

$$\text{(11)} \quad E(y_{t|x}) = \beta E(y_{t-1|x}) + E[g(w_t|x)].$$

Given a time-invariant covariate distribution $P(w|x)$, a time-invariant mean outcome $E(y_{x})$ determines a social equilibrium if $E(y_{x})$ solves the equation

$$\text{(12)} \quad E(y_{x}) = \beta E(y_{x}) + E[g(w_{x})].$$

Provided that $\beta < 1$, this equation has the unique solution

$$\text{(13)} \quad E(y_{x}) = E[g(w_{x})]/(1 - \beta).$$

Hence the unique equilibrium outcome distribution is
Dispersion-Dependent Strength of Norm: The strength of the effect of a social norm on individual behavior may depend on the dispersion of behavior in the reference group; the smaller the dispersion, the stronger the norm. A straightforward way to express this idea is to replace the parameter \( \beta \) in (10) with some function of the variance of outcomes in the reference group. Thus consider the model

\[
y_t = s[V(y_{t-1})_x] \cdot E(y_{t-1}_x) + g(w_t).
\]

Here \( s(\bullet): \mathbb{R}^+ \to \mathbb{R}^1 \) expresses the strength of the endogenous interaction as a function of the lagged group outcome variance \( V(y_{t-1}_x) \).

The laws of motion of the mean and variance of the group outcome distribution are

\[
E(y_{t|x}) = s[V(y_{t-1|x})] \cdot E(y_{t-1|x}) + E[g(w_t)_x]
\]

and

\[
V(y_{t|x}) = V[g(w_t)_x].
\]

Given a time-invariant covariate distribution \( P(w_x) \), a time-invariant mean outcome \( E(y_x) \) determines a social equilibrium if \( E(y_x) \) solves the equation

\[
E(y_x) = s[V[g(w)_x]] \cdot E(y_x) + E[g(w)_x].
\]
Thus, group x has a unique equilibrium if \( s[V[g(w)_x]]  \leq 1 \).

**Covariate-Dependent Strength of Norm:** The strength of a social interaction may vary across individuals. A straightforward way to express this idea is to replace the parameter \( \beta \) in (10) with some function of the individual covariates \( w_t \). Thus consider the model

\[
(19) \quad y_t = s(w_t) \cdot E(y_{t-1|_x}) + g(w_t).
\]

Here \( s(\bullet) : W \rightarrow R^1 \) expresses the manner in which the strength of the social interaction varies with the covariates \( w \).

The law of motion of the group mean outcome \( E(y_{t|x}) \) is

\[
(20) \quad E(y_{t|x}) = E[s(w_t)\cdot E(y_{t-1|x}) + E[g(w_t)]_x].
\]

Given a time-invariant covariate distribution \( P(w_x) \), a time-invariant mean outcome \( E(y_x) \) determines a social equilibrium \( x \) if \( E(y_x) \) solves the equation

\[
(21) \quad E(y_x) = E[s(w)_x] \cdot E(y_x) + E[g(w)_x].
\]

Thus, group x has a unique equilibrium if \( E[s(w)_x]  \leq 1 \).

**Linear-in-Location Model:** An obvious extension of the linear-in-mean model is to assume that \( y \) varies linearly with some location parameter \( L(\bullet) \) of the lagged group outcome distribution, not necessarily its mean. Thus let
(22) $y_t = \beta L(y_{t-1|x}) + g(w_t)$.

The location parameter $L(y_{t-1|x})$ could, for example, be the median or some other quantile of $P(y_{t-1|x})$. This model has the same dynamics as does the linear-in-mean model, with the location operator $L(\bullet)$ replacing the expectation operator $E(\bullet)$.

**Binary Response Models:** Models with nonlinear interactions are generally not as transparent as are the semi-linear models described above. Perhaps the most familiar nonlinear models are binary response models of the form

(23) $y_t = 1$ if $\beta P(y_{t-1} = 1|x) + g(w_t) > 0.$

$= 0$ otherwise.

As before, $\beta$ is a parameter expressing the strength of the social interaction and $g(\bullet): W \rightarrow R^1$ expresses the manner in which covariates influence outcomes.

The law of motion for $P(y_t = 1|x)$ is

(24) $P(y_t = 1|x) = P[w_t: \beta P(y_{t-1} = 1|x) + g(w_t) > 0|x]$.

Given a time-invariant covariate distribution, a time-invariant response probability $P(y = 1|x)$ is a social equilibrium in group $x$ if $P(y = 1|x)$ solves the equation

(25) $P(y = 1|x) = P[w: \beta P(y = 1|x) + g(w) > 0|x]$.

Manski (1993) shows that if the distribution function of $P[g(w)|x]$ is continuous and strictly increasing, then (25) has at least one solution. The equilibrium is unique if $\beta \leq 0$, but there may be multiple equilibria if $\beta > 0$. Brock
and Durlauf (1995) shows that if $P[g(w) x]$ has a logistic distribution and $\beta > 0$, then there is either a unique equilibrium or three equilibria, depending on the location and scale parameters of $P[g(w) x]$.

**Autoregressive Models:** Although our primary concern is with social interactions, it is worth noting that equation (7) encompasses autoregressive models of the form

\[(26) \quad y_t = f(y_{t-1}, w_t).\]

In this case, each person is his or her own reference group, with a distinct value of $x$. The distribution $P(y_{t-1} x)$ is degenerate with all its mass at the person’s realization of $y_{t-1}$.

3.2. Nonparametric Identification With Complete Data

The severity of the problem of inference on the response function $f(\cdot, \cdot)$ is determined by three factors:

1. the researcher’s maintained assumptions restricting the form of $f(\cdot, \cdot)$
2. the data available to the researcher
3. the values of $[P(y_{t-1} x), w_t]$ realized in the population.

There are a vast set of configurations of these factors that might usefully be explored. The researcher might assume nothing a priori about the form of the response function or might assume that $f(\cdot, \cdot)$ belongs to a tightly restricted family of functions. The data available to the researcher might be a sample of observations of some outcomes and covariates, or might be the entire distribution $P(x, w_t, y_t, t = 0, \ldots, T)$. The values of $[P(y_{t-1} x), w_t]$ realized in the population, termed the support of $[P(y_{t-1} x), w_t]$, might be a single point or the entire domain $G \times W$ of the response.
Although all three factors combine to determine the feasibility of inference on the response function, there is a fundamental difference between the first two factors and the third. Whereas the maintained assumptions and the available data characterize the researcher's contribution to inference, nature determines the support of \([P(y_{t-1,x}), w_t]\). The actual interaction process reveals the value of \(f(\bullet, \bullet)\) on the support but not at other points in \(G \times W\). Henceforth, the support of \([P(y_{t-1,x}), w_t]\) will be denoted \(\Omega\).

I focus here on the role of \(\Omega\) in inference when the researcher makes no assumptions restricting \(f(\bullet, \bullet)\), but observes the entire distribution \(P(x, w_t, y_t, t = 0, \ldots, T)\); a setting with unobserved covariates will be examined later. To simplify the exposition, I assume henceforth that the distribution \(P(x, w_t, t = 1, \ldots, T)\) of reference groups and covariates has discrete support. Let \(X_s\) denote the support of \(P(x)\); thus, \(x \in X_s\) if and only if group \(x\) contains a positive fraction of the population. For \(x \in X_s\), let \(W_{xt}\) denote the support of \(P(w_t_x)\). Let \(W_s = \{W_{xt}, t = 1, \ldots, T, x \in X_s\}\) denote the set of covariate values realized at some date by a positive fraction of the population. Then \(\Omega\) has the form

\[
\Omega = \{[P(y_{t-1,x}), w_t], w_t \in W_{xt}, t = 1, \ldots, T, x \in X_s\}.
\]

The shape of \(\Omega\) determines what can be learned nonparametrically about the response function, and, consequently, determines the extent to which inference must rely on assumptions restricting the form of \(f(\bullet, \bullet)\). Under the maintained assumption (7), the shape of \(\Omega\) is constrained by the law of motion (8). I have called this constraint on \(\Omega\) the reflection problem (Manski, 1993). The term reflection is appropriate here because, at each date \(t\), \(P(y_{t-1,x})\) is the aggregated image of the most recent outcomes of the members of group \(x\). The problem of identification of \(f(\bullet, \bullet)\)
is similar to the problem of interpreting the movements of a person standing before a mirror. Given data on the person's movements, one might ask what can be inferred about the way the person reacts to his most recent image in the mirror. Analogously, given data on the time path of outcomes in group $x$, we ask what can be inferred about the way the group members react at date $t$ to their most recent aggregated image $P(y_{t-1|_x})$. 

I want to focus on the shape of $\Omega$ when the population is observed in equilibrium. Suppose that in each group $x$, the covariate distribution $P(w_x)$ is time-invariant and the initial condition $P(y_{0|_x})$ is an equilibrium. Then $P(y_{t|_x}) = P(y_{0|_x})$ at each date $t$ and

$$\Omega = \{[P(y_{0|_x}), w], w \in W_x, x \in X_s\}.$$ 

With no time-series variation in covariates or interactions, inference on $f(\cdot, \cdot)$ must rely solely on cross-sectional variation.

When the population is observed in equilibrium, the shape of $\Omega$ depends on the unknown form of $f(\cdot, \cdot)$, the known distribution $P(x, w)$ of reference groups and covariates, and the known initial conditions $P(y_{0|_x})$. Let us focus on the distribution $P(x, w)$, which expresses how the population is sorted into reference groups. A simple but interesting finding is that, if $x$ and $w$ are either statistically independent or perfectly dependent, then $\Omega$ is a curve on $W_s$; that is,

$$\Omega = [(T_w, w), w \in W_s],$$

where $T_w \in G$ is a distribution that may vary with $w$.

Equation (29) means that, at each covariate value $w \in W_s$, $f(\cdot, w)$ is identified only at the single point $T_w$. Hence we can learn nothing about how $f(\cdot, w)$ varies
over G. See Proposition 3 of Manski (1993) for a proof under the assumption that interactions are transmitted through group-mean outcomes and that each group has a unique equilibrium. A more general proof applying to models of the form (7) is given in the Appendix to the present paper.

To illustrate, consider the determination of pregnancy among teenage girls. Let the pregnancy outcome $y_t$ of a girl in reference group $x$ be determined by the lagged pregnancy rate $P(y_{t-1} = 1 | x)$ of girls in the group and by the girl's covariates $w_t$. Let a girl's reference group $x$ be the neighborhood in which she lives. Let the relevant covariate $w_t$ be the girl's religion. Let the population of teenage girls be observed in pregnancy equilibrium.

In this setting, nonparametric study of the effect of endogenous interactions on pregnancy among girls of a given religion is impossible if every neighborhood has the same religious composition (statistical independence of $x$ and $w$), if neighborhoods are segregated by religion ($w$ perfectly dependent on $x$), or if all members of the given religion live in the same neighborhood ($x$ perfectly dependent on $w$). Each of these findings taken alone is simple, almost trivial. Yet the three findings taken together are intriguing. They imply that nonparametric inference is feasible only if neighborhoods and religions are "moderately associated" in the population.

3.3. Nonparametric Identification of The Reduced Form

Equation (7), which states the structural dependence of outcomes on group outcome distributions and individual covariates, implies the reduced-form

\[(30) \quad y_t = f(x, w_t),\]
where \( f(x, w_t) = f[P(y_{t-1|x}), w_t] \). It is important to distinguish the problem of identifying the response function \( f(\bullet, \bullet) \) from that of identifying the reduced-form functions \( f(\bullet, \bullet) \), \( t = 1, \ldots, T \). Whereas \( f(\bullet, \bullet) \) is nonparametrically identified on the set \( \Omega \) defined in (27), \( f(\bullet, \bullet) \) is nonparametrically identified on the set

\[
(31) \quad \Omega = \{(x, w_t), w_t \in W_{xt}, x \in X_s\}.
\]

If the support \( W_{xt} \) of \( P(w_{t|x}) \) is time-invariant for each \( x \in X_s \), then so is \( \Omega \), with

\[
(32) \quad \Omega^* = \{(x, w), w \in W_x, x \in X_s\}.
\]

Nonparametric study of the reduced-form variation of outcomes with reference groups is a simpler task than is nonparametric study of the structural variation of outcomes with reference-group outcome distributions. Suppose, for simplicity, that \( W_{xt} \) is time-invariant for each \( x \in X_s \), so (32) holds. We showed earlier that \( \Omega \) is a curve on \( W_s \) if \( x \) and \( w \) are either statistically independent or perfectly dependent. In contrast, \( \Omega^* \) is a curve on \( W_s \) only if \( x \) is perfectly dependent on \( w \).

Suppose that a researcher estimating the reduced form finds that, for each \( w \in W_s \) and \( t = 1, \ldots, T \), \( f(\bullet, w) \) does not vary with \( x \). This empirical finding does not suffice for the researcher to conclude that social interactions are absent. Constancy of \( f(\bullet, w) \) may reflect constancy of the outcome distributions \( P(y_{t-1|x}) \) across \( x \) and \( t \).

3.4. Identification of Response Functions Separable in Unobserved Covariates

There are many empirically relevant ways to weaken the idealized assumption that the researcher fully observes the population distribution \( P(x, w_t, y_t, \ldots) \).
t = 0,...,T) of reference groups, covariates, and outcomes. This section examines the problem of structural inference when some covariates are unobserved and the response function is separable in these unobserved covariates. Inference on separable response functions is a longstanding concern of the literature on econometric methods (see Manski, 1988, Sections 2.5 and 6.1).

Formally, let \( w_t = (z_t, u_t) \), where \( W = Z \times U \). Let the response function \( f(\cdot, \cdot) \) be known to have the separable form

\[
(33) \quad f[P(y_{t-1|x}), z_t, u_t] = h[P(y_{t-1|x}), z_t] + u_t,
\]

where \( h(\cdot, \cdot) : G \times Z \rightarrow Y \). The researcher observes the distribution \( P(x, z_t, y_t; t = 0,...,T) \) but does not observe \( u_t \).

Inference on \( f(\cdot, \cdot) \) is impossible in the absence of assumptions restricting the unobserved covariates. This obvious fact is often referred to as the "omitted-variables" problem. It is common in econometric studies to assume that the distribution of unobserved covariates conditional on observed covariates has zero mean, median, or other location parameter. In the present case, let

\[
(34) \quad L(u_t|x, z_t) = 0,
\]

where \( L \) denotes a specified location parameter. Taken together, assumptions (7), (33), and (34) are equivalent to assuming that at each date \( t = 1,...,T \),

\[
(35) \quad L(y_{t|x}, z_t) = h[P(y_{t-1|x}), z_t].
\]

Comparison of (7) and (35) shows that the problem of nonparametric identification of \( h(\cdot, \cdot) \) is precisely analogous to the problem of nonparametric identification of \( f(\cdot, \cdot) \) studied in Section 3.2. For each \( x \in X_s \), let \( Z_{xt} \) denotes the
support of $P(z_{t\mid x})$. Let $\Omega_h$ denote the support of $[P(y_{t-1\mid x}), z_t]$. That is, let

$$\Omega_h = \{[P(y_{t-1\mid x}), z_t], z_t \in Z_{xt}, t = 1,\ldots,T, x \in X_s\}.$$  

Observation of $P(x, z_t, y_t; t = 0,\ldots,T)$ reveals $h(\bullet, \bullet)$ on $\Omega_h$ but not elsewhere on the domain $G \times Z$.

Identification of $h(\bullet, \bullet)$ on $\Omega_h$ implies identification of the values of the covariates $u$. Thus assumption (35) overcomes entirely the identification problem generated by unobserved covariates. Consider a researcher who does not observe $u$ but knows $f(\bullet, \bullet)$ to be separable in $u$ and knows $u$ to have zero conditional location parameter. From the perspective of identification (although not from the perspective of statistical inference), this researcher is in the same position as one who fully observes reference groups, covariates, and outcomes.

4. Conclusion

This paper concludes with some questions answered. Analysis of the linear-in-means model in Section 2 shows that endogenous interactions, contextual interactions, and correlated effects are generally not distinguishable given observations of the outcomes and covariates of groups in equilibrium. These hypotheses are, however, distinguishable in principle (see note 3). The analysis of pure endogenous interactions models in Section 3 emphasizes that the law of motion of the interaction process constrains the data generated by this process, and so constrains the feasibility of nonparametric inference.

We also conclude with important issues unresolved. We have had almost nothing to say about inference when groups are observed out-of-equilibrium. The
time path of a process that is not in equilibrium is, by definition, more variable than the time path of a process that is in equilibrium. This suggests that observation of an out-of-equilibrium process generically enhances the feasibility of inference. What is missing thus far, however, is a constructive way to characterize the inferential benefits of out-of-equilibrium observations.

We have had too little to say about the way that the composition of reference groups affects the feasibility of inference. We have reported that nonparametric inference is impossible if $x$ and $w$ are statistically independent or perfectly dependent, but we have not characterized the possibilities for inference in the vast array of intermediate situations where $x$ and $w$ are moderately associated. We also have had too little to say about the vexing, and quite common, problem of inference when the researcher does not observe the composition of reference groups.

Some readers will naturally be interested in other models of social interactions: ones with time-varying reference groups or multiple reference groups, models with local rather than anonymous interactions, and so on. To the extent that such models constitute generalizations of the ones examined here, the negative identification findings reported here extend immediately. Our positive findings, however, do not necessarily extend to more general models.

Stepping further back from the particular concerns of this paper, I would reiterate my observation in the Introduction that the theoretical study of social interactions has long dominated the attention given to inference. Recent theoretical advances in our understanding of the economy as a complex evolving system, whether achieved by algebraic or computational means, are to be applauded. It is unfortunate that these advances have not been accompanied by corresponding progress in our grasp of the problem of empirical inference on actual economic processes.
Appendix

Proposition 1 gives the simple abstract reasoning underlying the finding, reported in Section 3.2, that nonparametric inference on endogenous interactions is infeasible if x and w are either statistically independent or perfectly dependent. Three Corollaries apply this reasoning to the cases of interest. The proofs are immediate.

**Proposition 1**: Let v \in W_s. Let X_v \perp (x \in X_s: v \in W_x). Let there exist a distribution Q_v on W such that P(w_t \mid x) = Q_v, t = 1,\ldots,t and x \in X_v. Let T_v \in G solve the social equilibrium equation

\[(A1) \quad T_v(y \in A) = Q_v[w: f(T_v, w) \in A], \quad \text{all } A \subseteq Y.\]

Let P(y_0 \mid x) = T_v, all x \in X_v. Then P(y_t \mid x) = T_v, t = 0,\ldots,t and x \in X_v. Hence f(\bullet, v) is identified only at T_v. •

**Corollary 1.1** (w statistically independent of x): Let there exist a distribution Q on W such that P(w_t \mid x) = Q, t = 1,\ldots,t and x \in X_s. Let T \in G solve the social equilibrium equation

\[(A2) \quad T(y \in A) = Q[w: f(T, w) \in A], \quad \text{all } A \subseteq Y.\]

Let P(y_0 \mid x) = T, all x \in X_s. Then P(y_t \mid x) = T, t = 0,\ldots,t and x \in X_s. Hence f(\bullet, w) is identified only at T, all w \in W_s. •

**Corollary 1.2** (w perfectly dependent on x): Let v \in W_s. Let X_v \perp (x \in X_s: v \in W_x). Let
almost all members of the groups $X_v$ have covariate value $v$; that is, let $P(w_t = v, x) = 1, t = 1, ..., t$ and $x \in X_v$. Let $T_v \in G$ solve the social equilibrium equation

\[(A3) \quad T_v(y \in A) = 1 \text{ if } f(T_v, v) \in A \]
\[= 0 \text{ otherwise, all } A \subseteq Y.\]

Let $P(y_0 | x) = T_v, x \in X_v$. Then $P(y_t | x) = T_v, t = 0, ..., t$ and $x \in X_v$. Hence $f(\cdot, v)$ is identified only at $T_v$. •

**Corollary 1.3** (x perfectly dependent on w): Let $v \in W_s$. Let almost all persons with covariate value $v$ belong to the same reference group, denoted $x(v)$; that is, let $P[x = x(v), w_t = v] = 1, t = 1, ..., T$. Let there exist a distribution $Q_v$ on $W$ such that $P[w_t | x(v)] = Q_v, t = 1, ..., t$. Let $T_v \in G$ solve the social equilibrium equation

\[(A4) \quad T_v(y \in A) = Q_v[w: f(T_v, w) \in A], \quad \text{all } A \subseteq Y.\]

Let $P[y_0 | x(v)] = T_v$. Then $P[y_t | x(v)] = T_v, t = 0, ..., t$. Hence $f(\cdot, v)$ is identified only at $T_v$. •
References


Schelling, Thomas (1971), "Dynamic Models of Segregation," Journal of
Mathematical Sociology, 1, 143-186.