Identification of Endogenous Social Effects: The Reflection Problem

CHARLES F. MANSKI

University of Wisconsin-Madison

This paper examines the reflection problem that arises when a researcher observing the distribution of behaviour in a population tries to infer whether the average behaviour in some group influences the behaviour of the individuals that comprise the group. It is found that inference is not possible unless the researcher has prior information specifying the composition of reference groups. If this information is available, the prospects for inference depend critically on the population relationship between the variables defining reference groups and those directly affecting outcomes. Inference is difficult to impossible if these variables are functionally dependent or are statistically independent. The prospects are better if the variables defining reference groups and those directly affecting outcomes are moderately related in the population.

1. INTRODUCTION

A variety of terms in common use connote endogenous social effects, wherein the propensity of an individual to behave in some way varies with the prevalence of that behaviour in some reference group containing the individual. These effects may, depending on the context, be called "social norms", "peer influences", "neighbourhood effects", "conformity", "imitation", "contagion", "epidemics", "bandwagons", "herd behaviour", "social interactions", or "interdependent preferences".

Endogenous effects have long been central to sociology and social psychology; see, for example, Asch (1952), Merton (1957), Erbring and Young (1979), and Bandura (1986). Mainstream economics has always been fundamentally concerned with a particular endogenous effect: how an individual's demand for a product varies with price, which is partly determined by aggregate demand in the relevant market. Economists have also studied other types of endogenous effects. Models of oligopoly posit reaction functions, wherein the output chosen by each firm is a function of aggregate industry output. Schelling (1971) analyzed the residential patterns that emerge when individuals choose not to live in neighbourhoods where the percentage of residents of their own race is below some threshold. Conlisk (1980) showed that, if decision making is costly, it may be optimal for individuals to imitate the behaviour of other persons who are better informed. Akerlof (1980), Jones (1984), and Bernheim (1991) studied the equilibria of non-cooperative games in which individuals are punished for deviation from group norms. Gaertner (1974), Pollak (1976), Alessie and Kapteyn (1991), and Case (1991) analyzed consumer demand models in which, holding price fixed, individual demand increases with the mean demand of a reference group.

While the price-mediated effect of aggregate demand on individual demand is quite generally accepted, other endogenous effects are controversial. Many economists regard

such central sociological concepts as norms and peer influences to be spurious phenomena explainable by processes operating entirely at the level of the individual. (See, for example, the Friedman (1957) criticism of Duesenberry (1949).) Even among sociologists, one still does not find consensus on the nature of social effects. For example, there has been a long-running debate about the existence and nature of neighbourhood effects. (See, for example, Jencks and Mayer (1989).)

Why do such different perspectives persist? Why do the social sciences seem unable to converge to common conclusions about the channels through which society affects the individual? I believe that a large part of the answer is the difficulty of the identification problem. Empirical analysis of behaviour often cannot distinguish among competing hypotheses about the nature of social effects.

Economists have long been concerned with the identification of endogenous effects channelled through markets, especially with the conditions under which observations of equilibrium prices and quantities reveal the demand behaviour of consumers and the supply behaviour of firms. But the identification of other endogenous effects has remained relatively unexamined and poorly understood.

This paper examines the "reflection" problem that arises when a researcher observing the distribution of behaviour in a population tries to infer whether the average behaviour in some group influences the behaviour of the individuals that comprise the group. The term reflection is appropriate because the problem is similar to that of interpreting the almost simultaneous movements of a person and his reflection in a mirror. Does the mirror image cause the person's movements or reflect them? An observer who does not understand something of optics and human behaviour would not be able to tell.

Although the reflection problem has several aspects, the series of simple findings reported in this paper collectively develop a theme. Inference on endogenous effects is not possible unless the researcher has prior information specifying the composition of reference groups. If this information is available, the prospects for inference depend critically on the population relationship between the variables defining reference groups and those directly affecting outcomes. Inference is difficult to impossible if these variables are functionally dependent or statistically independent. The prospects are better if the variables defining reference groups and those directly affecting outcomes are "moderately" related in the population.

Section 2 examines the reflection problem in the context of a linear model applied in many empirical studies of social effects. Section 3 analyzes non-linear models. Section 4 discusses dynamic models. Section 5 relates conventional consumer demand analysis to the work of this paper. Section 6 concludes by stressing the need for richer data if the analysis of social effects is to make more progress.

2. A LINEAR MODEL

The linear model analyzed here gives formal expression to three hypotheses often advanced to explain the common observation that individuals belonging to the same group tend to behave similarly. These hypotheses are:

(a) *endogenous effects*, wherein the propensity of an individual to behave in some way varies with the behaviour of the group;

(b) *exogenous (contextual) effects*, wherein the propensity of an individual to behave in some way varies with the exogenous characteristics of the group,[1] and

1. In the sociological literature, this is referred to as a "contextual effect". Inference on contextual effects became an important concern of sociologists in the 1960s, when substantial efforts were made to learn the effects on youth of school and neighbourhood environment (e.g. Coleman *et al.* (1966); Sewell and Armer (1966)). The recent resurgence of interest in spatial concepts of the underclass has spawned many new empirical studies (e.g. Crane (1991), Jencks and Mayer (1989), and Mayer (1991)). I use the term "exogenous" effect as a synonym for contextual effect, to distinguish the idea from endogenous effects.

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

An example may help to clarify the distinction. Consider the high school achievement of a teenage youth. There is an endogenous effect 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. There is an exogenous effect if achievement tends to vary with, say, the socio-economic composition of the reference group. There are correlated effects if youths in the same school tend to achieve similarly because they have similar family backgrounds or because they are taught by the same teachers.

The three hypotheses have differing policy implications. Consider, for example, an educational intervention providing tutoring to some of the students in a school but not to the others. If individual achievement increases with the average achievement of the students in the school, then an effective tutoring programme not only directly helps the tutored students but, as their achievement rises, indirectly helps all students in the school, with a feedback to further achievement gains by the tutored students. Exogenous effects and correlated effects do not generate this "social multiplier".

Section 2.1 specifies the model. Sections 2.2 and 2.3 analyse the identification problem, first considering the general model and then a restricted version assuming that neither exogenous nor correlated effects are present. Section 2.4 shows that, although the linear model sometimes imposes restrictions on observed behaviour, the model holds tautologically if the attributes defining reference groups and those directly affecting outcomes are functionally dependent. Section 2.5 draws implications for the problem of identifying reference groups. Section 2.6 discusses sample inference.

2.1. *Model specification*

Let each member of a population be characterized by a value for $(y, x, z, u) \in R^1 \times R^J \times R^K \times R^1$. Here y is a scalar outcome (e.g. a youth's achievement in high school), x are attributes characterizing an individual's reference group (e.g. a youth's school or ethnic group), and (z, u) are attributes that directly affect y (e.g. socioeconomic status and ability). A researcher observes a random sample of realizations of (y, x, z). Realizations of u are not observed.

Assume that

$$y = \alpha + \beta E(y|x) + E(z|x)'\gamma + z'\eta + u, \ E(u|x, z) = x'\delta, \tag{1}$$

where $(\alpha, \beta, \gamma, \delta, \eta)$ is a parameter vector. It follows that the mean regression of y on (x, z) has the linear form

$$E(y|x, z) = \alpha + \beta E(y|x) + E(z|x)'\gamma + x'\delta + z'\eta. \tag{2}$$

If $\beta \neq 0$, the linear regression (2) expresses an endogenous effect: a person's outcome y varies with $E(y|x)$, the mean of y among those persons in the reference group defined by x.[2] If $\gamma \neq 0$, the model expresses an exogenous effect: y varies with $E(z|x)$, the mean of the exogenous variables z among those persons in the reference group. If $\delta \neq 0$, the model expresses correlated effects: persons in reference group x tend to behave similarly

---

2. Beginning with Hyman (1942), sociological reference-group theory has sought to express the idea that individuals learn from or are otherwise influenced by the behaviour and attitudes of some reference group. Bank, Slavings, and Biddle (1990) give an historical account. Sociological writing has remained predominately verbal, but economists have interpreted reference groups as conditioning variables, in the manner of (2). See Alessie and Kapteyn (1991) or Manski (1993a).

because they have similar unobserved individual characteristics u or face similar institutional environments. The parameter η expresses the direct effect of z on y.[3]

2.2. *Identification of the parameters*

The question is whether the two types of social effects can be distinguished from one another and from the non-social effects. Thus, we are interested in identification of the parameter vector $(\alpha, \beta, \gamma, \delta, \eta)$. To focus attention on this question, I shall assume that either (i) (x, z) has discrete support or (ii) y and z have finite variances and the regressions $[E(y|x, z), E(y|x), E(z|x)]$ are continuous on the support of (x, z). Either assumption implies that $[E(y|x, z), E(y|x), E(z|x)]$ are consistently estimable on the support of (x, z).[4] So we can treat these regressions as known and focus on the parameters.

The reflection problem arises out of the presence of $E(y|x)$ as a regressor in (2). Integrating both sides of (2) with respect to z reveals that $E(y|x)$ solves the "social equilibrium" equation

$$E(y|x) = \alpha + \beta E(y|x) + E(z|x)'\gamma + x'\delta + E(z|x)'\eta. \tag{3}$$

Provided that $\beta \neq 1$, equation (3) has a unique solution, namely

$$E(y|x) = \alpha/(1-\beta) + E(z|x)'(\gamma+\eta)/(1-\beta) + x'\delta/(1-\beta). \tag{4}$$

Thus, $E(y|x)$ is a linear function of $[1, E(z|x), x]$, where 1 denotes the constant. It follows that the parameters $(\alpha, \beta, \gamma, \delta)$ are all unidentified. Endogenous effects cannot be distinguished from exogenous effects or from correlated effects.

What is identified? Inserting (4) into (2) we obtain the reduced form model

$$E(y|x, z) = \alpha/(1-\beta) + E(z|x)'[(\gamma+\beta\eta)/(1-\beta)] + x'\delta/(1-\beta) + z'\eta. \tag{5}$$

Inspection of (5) provides our first result:

Proposition 1. *In the linear model (2) with $\beta \neq 1$, the composite parameters $\alpha/(1-\beta)$, $(\gamma+\beta\eta)/(1-\beta)$, $\delta/(1-\beta)$, and η are identified if the regressors $[1, E(z|x), x, z]$ are linearly independent in the population.[5]*

Identification of the composite parameters does not enable one to distinguish between the two social effects but does permit one to determine whether some social effect is present. If $(\gamma+\beta\eta)/(1-\beta)$ is non-zero, then either $\beta\eta$ or γ must be non-zero.

3. Many generalizations of model (2) are of potential interest. Non-linear and dynamic models will be examined in Sections 3 and 4. Some other directions for generalization include the following:

    (i) Each person might be influenced by multiple reference groups, giving more weight to the behaviour of some groups than to others.

 (ii) The outcome y might be a vector, yielding a system of endogenous effects with mean reference-group outcomes along each dimension affecting individual outcomes along other dimensions.

(iii) Social effects might be transmitted by distributional features other than the mean. For example, it is sometimes said that the strength of the effect of social norms on individual behaviour depends on the dispersion of behaviour in the reference group; the smaller the dispersion, the stronger the norm.

4. Cell-average estimates may be used if the support of (x, z) is discrete. A variety of non-parametric regression estimates may be used if assumption (ii) holds; see, for example, Hardle (1990). Assumptions (i) and (ii) cover many but not all cases of empirical interest. They do not seem appropriate in studies of small-group social interactions, such as family interactions. In analyses of family interactions, each reference group (i.e. family) has negligible size relative to the population and random sampling of individuals typically does not yield multiple members of the same family. Hence, it is not a good empirical approximation to assume that (x, z) has discrete support. Moreover, unless one can somehow characterize different families as being similar in composition, one cannot assume that $[E(y|x, z), E(y|x), E(z|x)]$ are continuous. Random sampling of individuals is not an effective data-gathering process for the study of family interactions. It is preferable to use families as the sampling unit.

5. Linear independence in the population means that the support of the distribution of $[1, E(z|x), x, z]$ is not a proper linear subspace of $R^1 \times R^K \times R^J \times R^K$.

The ability to detect some social effect breaks down if $E(z|x)$ is a linear function of $[1, x, z]$. Unfortunately, $E(z|x)$ is a linear function of $[1, x, z]$ in various situations, including those stated in the following corollary to Proposition 1.

Corollary. *In the linear model* (2) *with* $\beta \neq 1$, *the composite social-effects parameter* $(\gamma + \beta\eta)/(1-\beta)$ *is not identified if any of these conditions hold, almost everywhere.*

(a) z *is a function of* x.
(b) $E(z|x)$ *does not vary with* x.
(c) $E(z|x)$ *is a linear function of* x.

The corollary shows that the ability to infer the presence of social effects depends critically on the manner in which z varies with x. In the context of the linear model (2), inference is possible only if $E(z|x)$ varies non-linearly with x and $\mathrm{Var}(z|x) > 0$.

2.3. *A pure endogenous-effects model*

The outlook for identification improves if one has information on some parameter values. Empirical studies of endogenous effects typically assume that $\gamma = \delta = 0$; so neither exogenous nor correlated effects are present. In this case, (2) reduces to

$$E(y|x, z) = \alpha/(1-\beta) + E(z|x)'[\beta\eta/(1-\beta)] + z'\eta. \qquad (6)$$

Inspection of (6) shows the following:

Proposition 2. *In the linear model* (2) *with parameter restrictions* $\gamma = \delta = 0$ *and* $\beta \neq 1$, *the composite parameters* $\alpha/(1-\beta)$, $\beta\eta/(1-\beta)$, *and* η *are identified if the regressors* $[1, E(z|x), z]$ *are linearly independent in the population.*

The endogenous-effects parameter β *is not identified if* $\eta = 0$ *or if* $E(z|x)$ *is a linear function of* $[1, z]$. *In particular,* β *is not identified if any of these conditions hold, almost everywhere*:

(a) z *is a function of* x.
(b) $E(z|x)$ *does not vary with* x.
(d) $E(z|x)$ *is a linear function of* x. x *is a linear function of* z.

For example, in a study of school achievement, β is identified if x is family income, z is ability, average ability $E(z|x)$ varies non-linearly with income, and achievement varies with ability (i.e. $\eta \neq 0$). But β is not identified if x is (ability, family income) and z is ability (condition a); if x is family income, z is ability, and average ability does not vary with income (condition b); or if x is family income, z is (ability, family income), and average ability varies linearly with income (condition d).

2.4. *Tautological models*

Even when its parameters are unidentified, a social-effects model may impose restrictions on observed behaviour and so have testable implications. There are, however, specifications of (x, z) that make a model hold tautologically. In particular, this is the case when x and z are functionally dependent.

Specify z to be a function of x, say $z = z(x)$. Then $E[y|x, z(x)] = E(y|x)$. So the linear model (2) holds with $\beta = 1$ and $\alpha = \gamma = \delta = \eta = 0$. Thus, observed behaviour is

always consistent with the hypothesis that individual behaviour reflects mean reference-group behaviour. For example, if a researcher studying student achievement specifies x to be (ability, family income) and z to be (ability), he will find that the data are consistent with the hypothesis that reference groups are defined by (ability, family income), that individual achievement reflects reference-group achievement, and that ability has no direct effect on achievement.

Conversely, specify x to be a function of z, say $x = x(z)$. Then $E[y|x(z), z] = E(y|z)$. So the semi-linear model

$$E(y|x, z) = \alpha + \beta E(y|x) + E(z|x)'\gamma + x'\delta + g(z) \tag{2'}$$

holds with $\alpha = \beta = \gamma = \delta = 0$ and $g(z) = E(y|z)$; the only testable restriction of the linear model (2) is its assumption that $g(\cdot)$ is a linear function. Continuing the school achievement example, a researcher who specifies x to be (family income) and z to be (ability, family income) will find that the data are consistent with the hypothesis that social forces do not affect achievement.

2.5. *Identifying reference groups*

So far, I have presumed that researchers know how individuals form reference groups and that individuals correctly perceive the mean outcomes experienced by their supposed reference groups. There is substantial reason to question these assumptions. Researchers studying social effects rarely offer empirical evidence to support their reference-group specifications. The prevailing practice is simply to assume that individuals are influenced by $E(y|x)$ and $E(z|x)$ for some specified x.[6] One of the few studies that does attempt to justify its specification of reference groups is Woittiez and Kapteyn (1991). They use individuals' responses to questions about their "social environments" as evidence on their reference groups.

If researchers do not know how individuals form reference groups and perceive reference-group outcomes, then it is reasonable to ask whether observed behavior can be used to infer these unknowns. The findings reported in Section 2.4 imply that this is not possible. Any specification of a functionally dependent pair (x, z) is consistent with observed behaviour. The conclusion to be drawn is that informed specification of reference groups is a necessary prelude to analysis of social effects.

2.6. *Sample inference*

Although our primary concern is with identification, a discussion of sample inference is warranted.

Empirical studies of social effects have generally assumed that there are no correlated effects and only one of the two types of social effects. Studies of exogenous effects have typically applied a two-stage method to estimate (γ, η). In the first stage, one uses the sample data on (z, x) to estimate $E(z|x)$ non-parametrically; typically x is discrete and the estimate of $E(z|x)$ is a cell-average. In the second stage, one estimates (γ, η) by finding the least squares fit of y to $[1, E_N(z|x), z]$, where $E_N(z|x)$ is the first-stage estimate of $E(z|x)$. See, for example, Coleman *et al.* (1966), Sewell and Armer (1966), Hauser (1970), Crane (1991) or Mayer (1991).

---

6. The same practice is found in empirical studies of decision making under uncertainty. Researchers assume they know how individuals form their expectations but offer no evidence justifying their assumptions. I have recently criticized this practice in the context of studies of schooling choice. See Manski (1993b).

Studies of endogenous effects have also applied a two-stage method to estimate (β, η), but in the guise of a "spatial correlation" model

$$y_i = \beta W_{iN} Y + z_i' \eta + u_i, \qquad i = 1, \cdots, N. \tag{7}$$

Here $Y = (y_i, i = 1, \cdots, N)$ is the $N \times 1$ vector of sample realizations of y and W_{iN} is a specified $1 \times N$ weighting vector; the components of W_{iN} are non-negative and sum to one. The disturbances u are usually assumed to be normally distributed, independent of x, and the model is estimated by maximum likelihood. See, for example, Cliff and Ord (1981), Doreian (1981), or Case (1991).

Equation (7) states that the behaviour of each person in the sample varies with a weighted average of the behaviours of the other sample members. Thus, the spatial correlation model assumes that an endogenous effect is present within the researcher's sample rather than within the population from which the sample was drawn. This makes sense in studies of small-group interactions, where the sample is composed of clusters of friends, co-workers, or household members; see, for example, Duncan, Haller, and Portes (1968) or Erbring and Young (1979). But it does not make sense in studies of neighbourhood and other large-group social effects, where the sample members are randomly chosen individuals. Taken at face value, equation (7) implies that the sample members know who each other are and choose their outcomes only after having been selected into the sample.

The spatial correlation model does make sense in studies of large-group interactions if interpreted as a two-stage method for estimating a pure endogenous-effects model. In the first stage, one uses the sample data on (y, x) to estimate $E(y|x)$ non-parametrically, and in the second stage, one estimates (β, η) by finding the least-squares fit of y to $[1, E_N(y|x), z]$, where $E_N(y|x)$ is the first-stage estimate of $E(y|x)$. Many non-parametric estimates of $E(y|x_i)$ are weighted averages of the form $E_N(y|x_i) = W_{iN} Y$, with W_{iN} determining the specific estimate; see Hardle (1990). Hence, estimates of (β, η) reported in the spatial correlation literature can be interpreted as estimates of pure endogenous-effects models.

Note that point estimates can be obtained for unidentified models. If condition a, b, or d of Proposition 2 holds, then $E(y|x)$ is a linear function of $[1, z]$. But the estimate $E_N(y|x)$ typically is linearly independent of $[1, z]$. So the two-stage procedure typically produces an estimate for β even when this parameter is unidentified.[7]

3. NON-LINEAR ENDOGENOUS-EFFECTS MODELS

How do the findings reported in Section 2 fare when the social-effects model is not necessarily linear? This section examines two situations. Section 3.1 assumes that one

---

7. It is necessary to point out that empirical studies reporting two-stage estimates of social-effects models have routinely misreported the sampling distribution of their estimates. The practice in two-stage estimation of exogenous-effects models has been to treat the first-stage estimate $E_N(z|x)$ as if it were $E(z|x)$ rather than an estimate thereof. The literature on spatial correlation models has presumed that equation (7) holds as stated and has not specified how the weights W_{iN} should change with N.

Two-stage estimation of social-effects models is similar to other semi-parametric two-stage estimation problems whose asymptotic properties have been studied recently. Ahn and Manski (1993), Ichimura and Lee (1991), and others have analyzed the asymptotic behaviour of various estimators whose first stage is non-parametric regression and whose second stage is parametric estimation conditional on the first-stage estimate. It is typically found that the second-stage estimate of an identified parameter is $\sqrt{N}$-consistent with a limiting normal distribution if the first-stage estimator is chosen appropriately. The variance of the limiting distribution is typically larger than that which would prevail if the first-stage regression were known rather than estimated. It seems likely that this result holds here as well.

does not know the form of the regression and so analyses social effects non-parametrically. Section 3.2 assumes that the regression is a member of a specified non-linear family of functions. To keep the analysis relatively simple, I restrict attention to pure endogenous-effects models.

3.1. *Non-parametric analysis*

Assume that, for some unknown function $f: R^1 \times R^K \to R^1$,

$$E(y|x, z) = f[E(y|x), z]. \tag{8}$$

This non-parametric endogenous-effects model, with the implied social equilibrium equation

$$E(y|x) = \int f[E(y|x), z] dP(z|x), \tag{9}$$

drops the linearity assumption imposed in Section 2.3.

In this non-parametric setting, one measures endogenous effects directly rather than through a parameter. That is, one fixes z and asks how $f[E(y|x), z]$ varies with $E(y|x)$. Let $\zeta \in R^K$ and $(e_0, e_1) \in R^2$. Then the contrast

$$T(e_1, e_0, \zeta) \equiv f(e_1, \zeta) - f(e_0, \zeta) \tag{10}$$

measures the effect at ζ of exogenously changing mean reference-group behaviour from e_0 to e_1. In the absence of functional form assumptions, $f(\cdot, \cdot)$ is identified on the support of $[E(y|x), z]$. The contrast $T(e_1, e_0, \zeta)$ is identified if and only if (e_1, ζ) and (e_0, ζ) are both on the support of $[E(y|x), z]$.

To say more requires that one characterize the support of $[E(y|x), z]$. Useful conditions ensuring that contrasts are identified seem hard to obtain. On the other hand, I can show that contrasts are generically not identified if x and z are either functionally dependent or statistically independent.

Proposition 3. *In the non-parametric endogenous-effects model* (8), *no contrasts of the form* (10) *are identified if any of these conditions hold, almost everywhere*:

  (e) *z is a function of x and the social equilibrium equation* (9) *has a unique solution.*
  (f) *z is statistically independent of x and the social equilibrium equation* (9) *has a unique solution.*
  (g) *x is a function of z.*

Proof. If $E(y|x)$ is a function of z, (e_1, ζ) and (e_0, ζ) cannot both be on the support of (x, z). So no contrasts are identified. Conditions e, f, and g all imply that $E(y|x)$ is a function of z.

  (e) Let $\zeta \in R^K$. The distribution of x conditional on the event $[z = \zeta]$ is concentrated on the set $X(\zeta) \equiv [x: z(x) = \zeta]$; hence, the distribution of $E(y|x)$ conditional on the event $[z = \zeta]$ is concentrated on $[E(y|x), x \in X(\zeta)]$. For $x \in X(\zeta)$, $P(z|x)$ has all its mass at the point ζ. Hence, for $x \in X(\zeta)$, equation (9) reduces to $E(y|x) = f[E(y|x), \zeta]$. So $E(y|x)$ solves the same equation for each x in $X(\zeta)$. The uniqueness assumption then implies that $E(y|x)$ is constant on $X(\zeta)$. So $E(y|x)$ is a function of z, almost everywhere.

(f) Statistical independence means that $P(z|x) = P(z)$. Hence (9) reduces to $E(y|x) = \int f[E(y|x), z]dP(z)$. So $E(y|x)$ solves the same equation for each value of x. The uniqueness assumption then implies that $E(y|x)$ is constant for all values of x.

(g) Let $x \equiv x(z)$. Then $E[y|x(z)]$ is a function of z.   ‖

A useful way to think about the proposition is to imagine that $f(\cdot, \cdot)$ is really linear with $\beta \neq 1$ but, not knowing this, one proceeds non-parametrically. The linear model has a unique social equilibrium so conditions e, f, and g all apply. Taken together, the conditions say that non-parametric identification of endogenous effects is precluded if the attributes defining reference groups and those directly affecting outcomes are functionally dependent or statistically independent. Non-parametric study of social effects remains conceivable only if x and z are "moderately related" random variables.

The prospects for identification may improve if $f(\cdot, \cdot)$ is non-linear in a manner that generates multiple social equilibria. In this case, condition g remains in effect but e and f do not apply. When there are multiple equilibria, $E(y|x)$ may fluctuate from one equilibrium value to another and so may not be a function of z.

3.2. Binary response models

Perhaps the most familiar non-linear parametric models with endogenous effects are binary response models. Let y be a binary random variable and assume that

$$P(y = 1|x, z) = H[\alpha + \beta P(y = 1|x) + z'\eta], \tag{11}$$

where $H(\cdot)$ is a specified continuous, strictly increasing distribution function. For example, if $H(\cdot)$ is the logistic distribution, we have a logit model with social effects.

Models of form (11) have been estimated by two-stage methods. The usual approach is to estimate $P(y = 1|x)$ non-parametrically and then estimate (β, γ) by maximizing the quasi-likelihood in which $P_N(y = 1|x)$ takes the place of $P(y = 1|x)$. Examples include Case and Katz (1991) and Gamoran and Mare (1989). A multinomial response model estimated in this manner appears in Manski and Wise (1983, Chapter 6).

The literature has not addressed the coherency and identification of model (11) but I can settle the coherency question here. The model is coherent if there is a solution to the social equilibrium equation

$$P(y = 1|x) = \int H[\alpha + \beta P(y = 1|x) + z'\eta]dP(z|x). \tag{12}$$

If $\beta = 0$, $E(y|x) = \int H(\alpha + z'\eta)dP(z|x)$ so there is a unique solution to (12). If $\beta < 0$, the right-hand side of (12) decreases strictly and continuously from $\int H(\alpha + z'\eta)dP(z|x)$ to $\int H(\alpha + \beta + z'\eta)dP(z|x)$ as $E(y|x)$ rises from 0 to 1. Meanwhile, the left-hand side increases strictly and continuously from 0 to 1. Hence the left- and right-hand sides cross at a unique value of $E(y|x)$.

Finally, let $\beta > 0$. In this case, a solution exists because the right-hand side of (12) increases strictly and continuously from $\int H(\alpha + z'\eta)dP(z|x)$ to $\int H(\alpha + \beta + z'\eta)dP(z|x)$ as $E(y|x)$ rises from 0 to 1. Meanwhile, the left-hand side traverses the larger interval $[0, 1]$. Hence, the left-hand side must cross the right-hand side from below at some value of $E(y|x)$.

The above shows that binary response models with endogenous effects are always coherent. When $\beta \leq 0$, these models have unique social equilibria. When $\beta > 0$, it does

not appear possible to determine the number of equilibria without imposing additional structure. The conditions under which the parameters (α, β, η) are identified have not been established.

4. DYNAMIC MODELS

The models posed thus far assume contemporaneous effects. It may well be more realistic to assume some lag in the transmission of these effects. Some authors, including Alessie and Kapteyn (1991) and Borjas (1991), have estimated the following dynamic version of the linear model (2):

$$E_t(y \mid x, z) = \alpha + \beta E_{t-1}(y \mid x) + E_{t-1}(z \mid x)' \gamma + x_t' \delta + z_t' \eta, \tag{13}$$

where E_t and E_{t-1} denote expectations taken at periods t and $t-1$. The idea is that non-social forces act contemporaneously but social forces act on the individual with a lag.

If $\{E(z \mid x), x, z\}$ are time-invariant and $-1 < \beta < 1$, the dynamic process (13) has a unique stable temporal equilibrium of the form (3). If one observes the process in temporal equilibrium, the identification analysis of Section 2 holds without modification. On the other hand, if one observes the process out of equilibrium, the recursive structure of (13) opens new possibilities for identification.

One should not, however, conclude that dynamic models solve the problem of identifying social effects. To exploit the recursive structure of (13), a researcher must maintain the hypothesis that the transmission of social effects really follows the assumed temporal pattern. But empirical studies typically provide no evidence for any particular timing. Some authors assume that individuals are influenced by the behaviour of their contemporaries, some assume a time lag of a few years, while others assume that social effects operate across generations.

5. DEMAND ANALYSIS

In Section 1, I noted that mainstream economic demand models embody an endogenous social effect: individual demand for a product varies with price, which is partly determined by aggregate demand in the relevant market. This section elaborates.

Let y denote a consumer's demand for a given product. Let x denote the market in which the consumer operates; different values of x may, for example, refer to different geographic areas or to different time periods. Let $p(x)$ be the market equilibrium price in market x. Then a conventional model of consumer demand assumes that, conditioning on consumer attributes, the market in which a consumer operates affects demand only through the price prevailing in that market. A common empirical formulation is

$$E(y \mid x, z) = D[p(x), z], \tag{14}$$

where z are consumer attributes observed by the researcher and where $D(\cdot, \cdot)$ is mean demand conditional on (x, z).

Market equilibrium models assume that the price $p(x)$ is determined by aggregate demand in market x and by supply conditions in this market. Let the population of consumers living in market x have size $m(x)$. Then $E(y \mid x)$ is per capita demand in market x and $E(y \mid x)m(x)$ is aggregate demand. Let $s(x)$ denote the relevant supply conditions. Then

$$p(x) = \pi[E(y \mid x)m(x), s(x)] \tag{15}$$

expresses the determination of price by demand and supply.

Equations (14) and (15) imply that

$$E(y|x, z) = D\{\pi[E(y|x)m(x), s(x)], z\}. \tag{16}$$

This is an endogenous effects model of a type distinct from (8). Conditional on z, y varies with x only through $E(y|x)$ in (8) but varies with x through $[E(y|x)m(x), s(x)]$ in (16). Equation (16) reduces to (8) if $m(\cdot)$ and $s(\cdot)$ do not vary with x; that is, if the population of consumers has the same size in all markets and if supply conditions are homogeneous across markets. In this case, the variation of price across markets derives entirely from variation in the distribution $p(z|x)$ of consumer attributes. The findings of Sections 2 and 3 then apply to the problem of identifying the consumer demand function.

6. CONCLUSION

This paper has analyzed the problem of identifying endogenous social effects from observations of the distribution of behaviour in a population. We have found that there may be realistic prospects for inference on endogenous effects if the attributes defining reference groups and those directly affecting outcomes are moderately related. On the other hand, the prospects are poor to nil if these attributes are either functionally dependent or are statistically independent. Moreover, observations of behaviour cannot be used to identify individuals' reference groups.

The only ways to improve the prospects for identification are to develop tighter theory or to collect richer data. I have no thoughts to offer on tighter theory but I see much that we can do to collect richer data. The analysis of this paper has presumed that inferences are based only on observed behaviour. Empirical evidence may also be obtained from controlled experiments and from subjective data, the statements people make about why they behave as they do. (Jones (1984) surveys some experiments conducted by social psychologists.) Given that identification based on observed behaviour alone is so tenuous, experimental and subjective data will have to play an important role in future efforts to learn about social effects.

REFERENCES

AHN, H. and MANSKI, C. (1993), "Distribution Theory for the Analysis of Binary Choice Under Uncertainty with Nonparametric Estimation of Expectations", *Journal of Econometrics* (forthcoming).
AKERLOF, G. (1980), "A Theory of Social Custom, of Which Unemployment may be One Consequence", *Quarterly Journal of Economics*, **94**, 749-775.
ALESSIE, R. and KAPTEYN, A. (1991), "Habit Formation, Interdependent Preferences and Demographic Effects in the Almost Ideal Demand System", *The Economic Journal*, **101**, 404-419.
ASCH, S. E. (1952) *Social Psychology* (Englewood Cliffs, N.J: Prentice-Hall).
BANDURA, A. (1986) *Social Foundations of Thought and Action* (Englewood Cliffs, N.J.: Prentice-Hall).
BANK, B., SLAVINGS, R. and BIDDLE, B. (1990), "Effects of Peer, Faculty, and Parental Influences on Students' Persistence", *Sociology of Education*, **63**, 208-225.
BERNHEIM, D. (1991), "A Theory of Conformity" (Department of Economics, Princeton University).
BORJAS, G. (1991), "Ethnic Capital and Intergenerational Mobility", *Quarterly Journal of Economics*, **108**, 123-150.
CASE, A. (1991), "Spatial Patterns in Household Demand", *Econometrica*, **59**, 953-965.
CASE, A. and KATZ, L. (1991), "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youth" (Working Paper No. 3705, National Bureau of Economic Research, Cambridge, Mass.).
CLIFF, A. and ORD, J. (1981) *Spatial Processes* (London: Pion).
COLEMAN, J., CAMPBELL, E., HOBSON, C., MCPARTLAND, J., MOOD, A., WEINFELD, F. and YORK, R. (1966) *Equality of Educational Opportunity* (Washington D.C.: U.S. Government Printing Office).

CONLISK, J. (1980), "Costly Optimizers Versus Cheap Imitators", *Journal of Economic Behavior and Organization*, **1**, 275-293.

CRANE, J. (1991), "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing", *American Journal of Sociology*, **96**, 1226-1259.

DUESENBERRY, J. (1949) *Income, Savings, and the Theory of Consumption* (Cambridge: Harvard University Press).

DUNCAN, O., HALLER, A. and PORTES, A. (1968), "Peer Influences on Aspirations: A Reinterpretation", *American Journal of Sociology*, **74**, 119-137.

ERBRING, L. and YOUNG, A. (1979), "Individuals and Social Structure: Contextual Effects as Endogenous Feedback", *Sociological Methods and Research*, **7**, 396-430.

FRIEDMAN, M. (1957) *A Theory of the Consumption Function* (Princeton: Princeton University Press).

GAERTNER, W. (1974), "A Dynamic Model of Interdependent Consumer Behavior", *Zeitschrift für Nationalökonomie*, **70**, 312-326.

GAMORAN, A. and MARE, R. (1989), "Secondary School Tracking and Educational Inequality: Compensation, Reinforcement, or Neutrality?", *American Journal of Sociology*, **94**, 1146-1183.

HARDLE, W. (1990) *Applied Nonparametric Regression* (New York: Cambridge University Press).

HAUSER, R. (1970), "Context and Consex: A Cautionary Tale", *American Journal of Sociology*, **75**, 645-664.

HYMAN, H. (1942), "The Psychology of Status", *Archives of Psychology*, No. 269.

ICHIMURA, H. and LEE, L. (1991), "Semiparametric Estimation of Multiple Indices Models: Single Equation Estimation", in W. Barnett, J. Powell, and G. Tauchen (editors), *Nonparametric and Semiparametric Methods in Econometrics and Statistics* (New York: Cambridge University Press).

JENCKS, C. and MAYER, S. (1989), "Growing Up in Poor Neighborhoods: How Much Does it Matter?", *Science*, **243**, 1441-1445.

JONES, S. (1984) *The Economics of Conformism* (Oxford: Basil Blackwell).

MANSKI, C. (1993a), "Dynamic Choice in a Social Setting: Learning from the Experiences of Others", *Journal of Econometrics*, **58**, 121-136.

MANSKI, C. (1993b), "Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?", in C. Clotfelter and M. Rothschild (editors), *Studies in the Supply and Demand of Higher Education* (Chicago: University of Chicago Press).

MANSKI, C. and WISE, D. (1983) *College Choice in America* (Cambridge; Mass.: Harvard University Press).

MAYER, S. (1991), "How Much Does a High School's Racial and Socioeconomic Mix Affect Graduation and Teenage Fertility Rates?", in C. Jencks and P. Peterson (editors), *The Urban Underclass* (Washington, D.C.: The Brookings Institution).

MERTON, R. (1957) *Social Theory and Social Structure* (Glencoe: The Free Press).

POLLAK, R. (1976), "Interdependent Preferences", *American Economic Review*, **78**, 745-763.

SCHELLING, T. (1971), "Dynamic Models of Segregation", *Journal of Mathematical Sociology*, **1**, 143-186.

SEWELL, W. and ARMER, J. (1966), "Neighborhood Context and College Plans", *American Sociological Review*, **31**, 159-168.

WOITTIEZ, I. and KAPTEYN, A. (1991), "Social Interactions and Habit Formation in a Labor Supply Model" (Department of Economics, University of Leiden).