A Framework for the Study of Individual Behavior and Social Interactions

Steven N. Durlauf
Department of Economics, University of Wisconsin at Madison

August 27, 2000

I thank the John D. and Catherine T. MacArthur Foundation, National Science Foundation, Vilas Trust and Romnes Trust for financial support. Thanks go to Aimée Dechter for many helpful conversations and to Mark Becker and Michael Sobel for initiating this paper. Special thanks go to William Brock; much of this paper describes our joint work and reflects our ongoing collaboration. Artur Minkin and Eldar Nigmatullin have provided excellent research assistance.

A Framework for the Study of Individual Behavior and Social Interactions

Abstract

Recent work in economics has begun to integrate sociological ideas into the modelling of individual behavior. In particular, this new approach emphasizes how social context and social interdependences influence the ways in which individuals make choices. This paper provides an overview of an approach to integrating theoretical and empirical analysis of such environments. The analysis is based on a framework due to Brock and Durlauf (2000a,2000b). Empirical evidence on behalf of this perspective is assessed and some policy implications are explored.

Steven N. Durlauf
Department of Economics
University of Wisconsin
1180 Observatory Drive
Madison, WI 53706-1393
sdurlauf@ssc.wisc.edu

1. Introduction

"...just as our political life is free and open, so is our day-to-day life in our relations with each other. We do not get into a state with our next-door neighbour if he enjoys himself in his own way, nor do we give him the type of black looks which, though they do no real harm, still do hurt people's feelings. We are free and tolerant in our private lives; but in public affairs we keep to the law. This is because it commands our deep respect. We give our obedience to those whom we put in positions of authority and we obey the laws themselves, especially those which are for the protection of the oppressed, and those unwritten laws which it is an acknowledged shame to break."

Pericles' Funeral Oration to the Athenians, c. 431-430 BCE
Thucydides, History of the Peloponnesian War (2.37)

"...the Athenians owed to the plague the beginnings of a state of unprecedented lawlessness. Seeing how quick and abrupt were the changes of fortune which came to the rich who died and to those who had previously been penniless but now inherited their wealth, people now began openly to venture on acts of self-indulgence which before they used to keep dark...As for what is called honour, no one showed himself willing to abide by its laws, so doubtful was it that one would survive to enjoy the name for it. It was generally agreed that what was both honourable and valuable was the pleasure of the moment."

Description of Effects of Plague in Athens, 430 BCE
Thucydides, History of the Peloponnesian War (2.53)[1]

This paper is designed to describe an approach for integrating social interactions into economic models. While social interactions, broadly defined to include phenomena ranging from societal norms to role models to networks, are a fundamental part of historical studies and other social sciences, especially sociology of course, they have only recently begun to play a prominent role in economic thinking. In this new work, the explicit choice-based reasoning of economic theory is extended to environments in which direct interdependences between socioeconomic actors (i.e. interdependences which are not mediated by markets) are of primary importance. This new research is explicitly designed to extend the domain of inquiry by economists into areas which have been the traditional domain of other social scientists. This new work does not, however, make this attempt out of dissatisfaction with these social sciences. Rather, the objective of this research program is to internalize within formal

[1]Taken from R. Walters translation, New York: Penguin Books, 1978.

economic models a number of the substantive ideas and perspectives of these other disciplines.

One area where interactions play a primary role is the new growth economics. Neoclassical growth models, stemming from Solow's seminal work (1957), have usually focused on the role of physical and human capital accumulation on producing growth. Technological change in this framework is treated as exogenous, i.e. something outside the domain of the analysis. The new approach to growth, launched by Lucas (1988) and Romer (1986), in contrast, explicitly models technical change and productivity growth as the outcomes of interactions between workers and between firms, and hence is known as "endogenous" growth theory.

To see how the analysis of technical change is facilitated by an emphasis on interactions, consider the question of the effect of an innovation by one firm on others. As has been emphasized in the new growth literature, innovations are at least partially nonrival, which means that the use of an innovation by one economic actor does not effect the ability of another to use it, and nonappropriable, which means that an innovator cannot fully exclude others from exploiting his advance. Together, these features mean that there are powerful externalities associated with the creation of new ideas. An innovation by one firm alters the production opportunities of other firms in ways which cannot be captured through patents. (A good example is the graphical user interface; Apple's initial idea affected Microsoft's technological development). When these spillover effects are strong enough, the aggregate technology frontier will exhibit increasing returns to scale, so that innovations and growth by one firm expand the innovations and growth of others creating large feedback effects. These types of spillovers may also possess a spatial dimension; Arthur (1987) uses them to explain the emergence of the Silicon Valley and other technology regional centers.

Interactions-based thinking has assumed a particularly prominent role in recent studies of poverty and inequality. In one facet of this new approach, the impact of residential neighborhoods on the future prospects of children has been explored. Bénabou (1993,1996a,b) and Durlauf (1996a,b) construct models of persistent intergenerational inequality based on the presence of spillover effects from the educational and economic characteristics of a neighborhood on human capital acquisition by children. In these types of models, children are assumed to be influenced by a range of community characteristics. One source of community influences is institutional – because of local public finance of education, the affluence of a community affects the level of per capita spending on schools.[2] Another source of community effects occurs via role models. If individual aspirations and assessment of educational effort depend on the observed education levels and associated occupations of adults in a community, then stratification of communities by income and education will induce cross-community differences in the educational efforts and attainment of children.

In another area of the new inequality research, social interactions have been identified as influencing labor market prospects via information flows (Montgomery (1992)). In Montgomery's framework, successful labor market matches are facilitated by social connections, which communicate job opportunities among members of a network. When communities exhibit disparate unemployment rates, differential access to new jobs will exist, thereby perpetuating the initial disparities; evidence consistent with this phenomenon has been found by Topa (1999).

The recent attention on interactions and inequality resonates with a number of interesting ethnographic studies in sociology. Anderson (1999), for example, provides the following description of the sources of inner city violence:

"The inclination to violence springs from the circumstances of life among the ghetto poor – the lack of jobs that pay a living wage, limited basic public services (police response in emergencies, building maintenance, trash pickup, lighting...), the stigma of

[2]There is considerable controversy concerning the effects of educational expenditure on educational outcomes. Hanushek (1996) has argued that this type of effect is negligible, when test scores are the outcome variable of interest. In contrast, Card and Krueger (1992) find that predictions of future wages are sensitive to educational quality. I do not take a strong stand on this question, except to say that the empirical literature has typically focused on linear models, whereas the effects may be nonlinear. Certainly Kozol (1991) is consistent with this view, in the sense that he documents how very poor schools are handicapped in the education they provide. One reason for this type of nonlinearity is that while schools may differ widely in the efficiency with which they use revenues, some minimum is needed for each educational quality level. Of course, nonlinearities may also imply that the Card and Krueger results are questionable. For example, Heckman, Layne-Farrar, and Todd (1996) find that the effects of school resources on labor market outcomes vary widely according to what control variables are included and according to educational group. For example, it appears that school quality matters primarily for workers who end up going to college. My own view of this literature is that while some quality effects are present, little is known about the causal mechanisms or even functional forms.

race, the fallout from rampant drug use and drug trafficking, and the resulting alienation and absence of hope for the future. Simply living in such an environment places young people at special risk of falling victim to aggressive behavior. Although there are often forces in the community that can counteract the negative influences – by far the most important is a strong loving...family that is committed to middle class values – the despair is pervasive enough to have spawned an oppositional culture, that of "the street" as consciously opposed to those of mainstream society. (pg. 22-23)

Thick descriptions of the type found in Anderson or Duneier (1992,1999) are an important reminder of the limits of the type of formal analysis which is used in this paper. Formal modelling of the type I describe, in many respects, only crudely approximates the many subtleties which are associated with social interactions; phenomena related to personal identity which Anderson explores are a good example.[3] On the other hand, to the extent that the objective of a research program is the construction of predictive or evaluative distributions of the effects of alternative policies, then the sort of formalization I describe is essential. To take a classic example, the Coleman report on the determinants of educational outcomes, was based upon and ultimately discredited through formal analysis. Alternatively, the most important work on the effects of Head Start and other social programs on individual outcomes has proceeded from the sort of quantitative social science I describe.[4]

2. A formal model framework for individual choice

i. basic ideas

In order to see how a social interactions perspective may be incorporated into standard economic reasoning, it is useful to start with a baseline description of individual choices for a population of individuals indexed by $i = 1 \ldots I$. Economists typically assume that the choice of individual i, ω_i, is interpretable as maximizing some payoff function V subject to a set Ω_i of possible choices which are available to

[3]Akerlof and Kranton (1999) is an important recent effort at grappling with these issues.

[4]See Currie and Thomas (1995) for a recent example.

that individual. With respect to the function V, whose form embodies individual preferences, the primitive modelling assumption is actually *not* that individuals literally possess these functions and explicitly calculate payoffs from alternative courses of action by using them. Rather the primitive notion is that individuals possess preference orderings over the space of possible choices they face; when these preference orderings fulfill certain axioms they may be mathematically represented by a payoff function.

The choice ω_i can therefore be treated as the solution to a maximization problem. For my purposes, this optimization problem possesses the generic form

$$\omega_i = argmax_{\omega \in \Omega_i} V(\omega, Z_i, \epsilon_i). \tag{1}$$

This particular payoff function is expressed as containing two distinct arguments beyond the choice ω_i: a vector of individual-indexed characteristics Z_i, which allows for observable heterogeneity in how individuals evaluate choices, and a vector of individual-indexed characteristics ϵ_i, which are assumed to be unobservable to a modeller, but are known to individual i, thereby allowing for unobservable heterogeneity. Introducing unobservable as well as observable heterogeneity is important in developing the theory in a direction which permits empirical implementation. This distinction is precisely the same as that between regressors and the disturbance in the specification of a linear regression.

The incorporation of social interactions into this framework is, at one level, nothing more than a particular choice as to what variables to include in Z_i. Suppose that individuals are associated with groups in a way that $g(i)$ identifies individual i's group. For purposes of developing the basic theory, I will assume that all members of the population are members of the same group; this assumption is relaxed when I consider empirical work.[5] Interactions can then be incorporated into individual decisions by including variables which depend on i only through information which is measured at the level of $g(i)$, i.e. interactions are modelled as the dependence of individual payoffs on group-indexed variables. The average education level among parents in a community or the average rate of cigarette smoking among teenagers in a

[5]It is relatively straightforward to generalize the theoretical analysis to the case where individuals are members of different groups.

given ethnic group are examples of such variables.

In developing the econometric version of a choice-based framework, it is important to distinguish between variables which represent the influence a group's characteristics have on its members and those variables which represent the influence a group's joint behaviors have on its members. Following Manski (1993), variables which measure the former represent *contextual* effects whereas variables which measure the latter are *endogenous* effects. In the context of youth behavior, role models constitute contextual effects whereas the contemporaneous behaviors of friends constitute endogenous effects. This language, of course, closely parallels usage from sociology, cf. Blalock (1984). This consideration means that it is convenient to separate Z_i into three distinct components: X_i, which represents a vector of variables which can vary across individuals within the group $g(i)$, $Y_{g(i)}$, which represents a vector of variables which are common to all members of the same group and are predetermined with respect to group behavior,[6] and $\mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i)$, which represents the beliefs of individual i concerning the choices of the other members of his group, given his information set F_i. The associated decision problem and choice of individual i can therefore be rewritten as:

$$\omega_i = argmax_{\omega \in \Omega_i} V(\omega, X_i, Y_{g(i)}, \mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i), \epsilon_i) \tag{2}$$

The decision to treat endogenous effects as occurring through beliefs concerning the behavior of others rather than their actual behavior (i.e. the choice of $\mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i)$ rather than $\underset{\sim}{\omega}_{-i}$ as the endogenous component) makes the theoretical analysis of the model substantially simpler.[7] The appropriateness of this assumption, of course, will depend on the particular context in which the model is employed; one would think, for example, if an individual cares about the aggregate characteristics of a large group, then the expectations assumption makes particular sense.

In order to close the model, it is necessary to specify how expectations are

[6]In the theoretical model I develop, $Y_{g(i)}$ is constant across i since all members of the population form a single group, and so its inclusion might seem redundant. I introduce this class of variables here in order to keep the theoretical model and its econometric implementation as close as possible.

[7]See Glaeser and Scheinkman (2000a,b) for analyses which use realized behaviors in the payoff function.

formed. Within economics, the standard assumption is that expectations are rational, which means that the subjective beliefs of individuals are consistent with the conditional probabilities which actually characterize the variables over which these beliefs are formed. Operationally, rational expectations may be thought of as follows. Suppose that each individual choice solves an optimization problem as described by eq. (2). This decision problem implies that the choice of each individual can be represented as

$$\omega_i = m(X_i, Y_{g(i)}, \mu_i^e(\underset{\sim}{\omega} \mid F_i), \epsilon_i) \; i = 1 \ldots I \tag{3}$$

for some choice function m. Here I have replaced $\mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i)$ with $\mu_i^e(\underset{\sim}{\omega} \mid F_i)$; one can always do this by adding the variable $E_i(\omega_i)$ to the payoff function V and assuming that different values of the variable leave V unaffected. Stacking these I choice functions together, one has a vector function R (whose elements correspond to the m-functions for each individual) such that

$$\underset{\sim}{\omega} = M(X_1, \ldots, X_I, Y_{g(1)}, \ldots, Y_{g(I)}, \mu_1^e(\underset{\sim}{\omega} \mid F_1), \ldots, \mu_I^e(\underset{\sim}{\omega} \mid F_I), \epsilon_1 \ldots \epsilon_I) \tag{4}$$

Next assume that all individuals possess identical information sets from which they form expectations about the choices in the population. Denote this common information set as F. Further, assume that this information set consists of the values of X_i and $Y_{g(i)}$ for all i in the population. By construction, each individual uses this information to form identical expectations $\mu^e(\underset{\sim}{\omega} \mid F)$. One can, using eq. (4), compute the actual conditional probability of the vector of choices given the model and the information set F. From eq. (4), it follows that

$$\mu(\underset{\sim}{\omega} \mid F) = \mu(\underset{\sim}{\omega} \mid F, \mu^e(\underset{\sim}{\omega} \mid F)). \tag{5}$$

If there exists a probability measure $\mu(\underset{\sim}{\omega} \mid F)$ such that

$$\mu(\underset{\sim}{\omega} \mid F) = \mu(\underset{\sim}{\omega} \mid F, \mu(\underset{\sim}{\omega} \mid F)) \tag{6}$$

then this measure $\mu(\underset{\sim}{\omega} \mid F)$ is a rational expectations solution to the model. In words, agents possess rational expectations conditional on F when their beliefs given F, as represented by conditional probabilities associated with an environment, are confirmed by the actual conditional probabilities generated by that environment. Notice that all I have done is define what rational expectations means; nothing has been said about the conditions under which a set of self-consistent beliefs exists, or if it exists, whether it is unique. Properties such as these can only be assessed in the context of particular specifications of an environment.

The rational expectations assumption is controversial, and dissatisfaction with it has led, over the last 15 years, to a rich literature on bounded rationality (Rubinstein (1998) provides a profound overview of this work). However, the literature on interactions-based models has generally not incorporated this approach in an interesting fashion,[8] and I will assume rationality in what follows. One should note that the rationality assumption is, in certain respects, inessential for understanding the qualitative properties of these systems. For example, the rational expectations equilibria of a static model often prove to be the limit points of various learning schemes (see Brock and Durlauf (1998,2000a) and Glaeser and Scheinkman (2000b) for specific examples of this). Further, the interesting qualitative properties of interactions-based models do not rely on rational expectations per se but rather are generated by the presence of feedbacks between group and individual behaviors.

Notice that this abstract description of behaviors in a population under rational expectations incorporates very standard economic reasoning. Individual decisions are explicitly modelled as purposeful choices, and a consistency condition is imposed across

[8]In saying that bounded rationality or learning spillovers have not been dealt with in an "interesting" way, I mean that many if not most models which claim to be based on bounded rationality fail to generate insights which differ from models where agents interact through preferences. To be a bit more precise, one can model the effects of the behavior of peers on an individual as due to two distinct factors: 1) a psychological desire to conform to one's peers, which is a claim about interdependent preferences, and 2) an information effect whereby the behavior of others is used by an individual to determine which choice is better for him, which (when not formulated as an optimal extraction of information) is a form of bounded rationality. On the other hand, there is an important related literature on social learning in which the behavior of individuals alters the information sets of others in an environment in which each individual acts rationally conditional on a limited information set; see Bikhchandani, Hirshleifer, and Welch (1992) for a very nice analysis of this type.

the choices, in this case consistency of beliefs with the probabilistic structure of the population's behavior. This combination of individual maximization and self-consistency is no different from what occurs when one specifies a set of individual demand and supply functions for a group of commodities, each of which takes prices as given, and then requires that prices clear markets.

3. Binary choice with social interactions

i. general formulation

The general choice framework I have described cannot be analyzed in more detail without placing greater structure on the choice problem. For example, one cannot say anything about the existence of multiplicity of equilibria without moving to a more specific formulation. One common approach to adding such structure is based on modelling binary choices, which is the case where $\Omega_i = \{-1,1\}\ \forall\ i$. For many of the areas of substantive applications of models of choice with social interactions, the binary assumption is a reasonable first-order approximation; specific examples include decisions on whether to use drugs, drop out of school, commit a crime, have a child out-of-wedlock etc. An additional advantage of focusing on binary choice problems is that many of the key theoretical ideas in the interactions literature can be easily understood in this context. For these reasons, I will use a binary choice framework developed in Brock and Durlauf (2000a,b) for the subsequent analysis in this paper. In the subsequent development, I will sequentially add more structure to the binary choice problems faced by population members in order to produce sequentially more precise descriptions of the properties of the resultant environment. For binary choices, the individual decision process can be expressed as

$$\omega_i = argmax_{\omega \in \{-1,1\}} V(\omega, X_i, Y_{g(i)}, \mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i), \epsilon_i) \tag{7}$$

Substantive insights into the role of interdependences between choices on aggregate behavior can be developed once some assumptions are placed on the structure of the

$V(\cdot,\cdot,\cdot,\cdot,\cdot)$ function. These assumptions will also render the model falsifiable.

Following Brock and Durlauf (2000a,b), the first assumption is that the general payoff function is additively separable,

$$V(\omega_i, X_i, Y_{g(i)}, \mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i), \epsilon_i) =$$

$$u(\omega_i, X_i, Y_{g(i)}) + S(\omega_i, X_i, Y_{g(i)}, \mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i)) + \epsilon_i(\omega_i) \tag{8}$$

Here, $u(\omega_i, X_i, Y_{g(i)})$ denotes private deterministic utility, $S(\omega_i, X_i, Y_{g(i)}, \mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i))$ denotes social deterministic utility and $\epsilon_i(\omega_i)$ denotes a private random (from the perspective of the modeller) utility. Notice that this term is now made an explicit function of ω_i. What this means is that the two choices may have differential effects on the payoff function. So, for example, if one is choosing between a career as a musician ($\omega_i = 1$) or as a painter ($\omega_i = -1$), $\epsilon_i(1)$ represents unobservable musical talent and $\epsilon_i(-1)$ represents unobservable artistic talent.

The additive separability assumption is made for two reasons. First, separating out the random term $\epsilon_i(\omega_i)$ is essential in achieving analytic tractability for the model. Second, this formulation is attractive in terms of empirical implementation. It will turn out that when $S(\omega_i, X_i, Y_{g(i)}, \mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i)) = 0$, the individual behavioral rule reduces to that described by the standard binary choice model, hence one will be able to test for social interactions.

Second, social utility is assumed to possess a particular functional form,

$$S(\omega_i, X_i, Y_{g(i)}, \mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i)) = -E_i\left(\sum_{j \neq i} \frac{J_{i,j}}{2} (\omega_i - \omega_j)^2\right) \tag{9}$$

where E_i refers to the subjective expected value assessment made by i given information F_i. The $J_{i,j}$ terms represent weights which apply to the bilateral interactions between members of the population. At this level of generality, one can allow for each of these values to differ. The quadratic functional form is designed to capture conformity effects. If $J_{i,j} > 0$, this functional form implies that individual i derives higher utility (other things being equal) from making the same choice as individual j; $J_{i,j} < 0$ in turn implies a utility benefit from acting differently. By choosing different distributions for the $J_{i,j}$'s, one can, in principle, model populations where incentives to conform and deviate coexist. One example of this may be dialect use, where the choice of nonstandard forms of grammar and syntax appear to stem from a desire for membership in some groups and a rejection of identification with others.[9]

Finally, one completes the model by choosing a distribution for the random terms $\epsilon_i(\omega_i)$. This is done by assuming that the difference in the random utility terms is logistically distributed,

$$Prob(\epsilon_i(-1) - \epsilon_i(1) \leq z) = \frac{1}{1+exp(-\beta_i z)};\ \ \beta_i \geq 0. \tag{10}$$

As in the case of the other assumptions, this functional form provides benefits in terms of analytics as well as a way of linking the theoretical model to a statistical one. Notice that β_i indexes the support of the unobserved heterogeneity. Roughly speaking, the larger the value of β_i, the less likely are large draws of $|\epsilon_i(-\omega_i) - \epsilon_i(\omega_i)|$.

Under these assumptions, it is possible to derive some expressions which describe the set of population choices. Before doing so, there is a simplification which can be made in the fairly abstract representation which I have described. First, one can replace the general utility function $u(\omega_i, X_i, Y_{g(i)})$ with a linear function

$$u(\omega_i, X_i, Y_{g(i)}) = h_i\omega_i + k_i \tag{11}$$

where the slope term h_i and intercept term k_i are chosen so that

$$h_i + k_i = u(1, X_i, Y_{g(i)}) \tag{12}$$

and

[9]For example, as described in Chambers (1995), the dropping of the letter g in words ending in $-ing$ is strongly associated with being poor and male in the US, UK and Australia. This is generally explained by the need for poorer males to develop an identity which rejects conventional metrics of success. The framework described here would seem to be a natural way of exploring the use of African American Vernacular English versus standard dialects.

$$-h_i+k_i=u(-1,X_i,Y_{g(i)}) \tag{13}$$

This simplification is allowable because choices are binary; so long as the new linear utility function matches the original V function when ω_i equals either -1 or 1, (which is what eqs. (12) and (13) impose via the implied restrictions on h_i and k_i) it is irrelevant that it fails to match the original function for other values of ω_i.

The model now has enough detail to allow a parametric description of the conditional probabilities of the vector of choices $\underset{\sim}{\omega}$. One does this first by calculating

$$\mu(\omega_i \mid X_i,Y_{g(i)},\mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i)) \tag{14}$$

the conditional probability of individual i's choice given his observable characteristics and his beliefs. Since the choice ω_i is only made when the payoff from the choice exceeds that which would be generated by $-\omega_i$, for any information set the conditional probability of ω_i is equal to the probability that the payoff at ω_i is greater than the payoff at $-\omega_i$, i.e.

$$\mu(\omega_i \mid X_i,Y_{g(i)}\ \mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i)) =$$

$$\mu(V(\omega_i,X_i,Y_{g(i)},\mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i),\epsilon_i(\omega_i)) > V(-\omega_i,X_i,Y_{g(i)},\mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i),\epsilon_i(-\omega_i))). \tag{15}$$

Substituting in the additively separable representation of the payoff function, eq. (8), and the linearized deterministic private utility function, eq. (11), this inequality may be rewritten as

$$\mu(h_i\omega_i - E_i\sum_{j\neq i}\frac{J_{i,j}}{2}(\omega_i-\omega_j)^2+\epsilon_i(\omega_i) > -h_i\omega_i - E_i\sum_{j\neq i}\frac{J_{i,j}}{2}(-\omega_i-\omega_j)^2+\epsilon_i(-\omega_i)) =$$

$$\mu(h_i\omega_i+\sum_{j\neq i}J_{i,j}\omega_iE_i(\omega_j)+\epsilon_i(\omega_i) > -h_i\omega_i-\sum_{j\neq i}J_{i,j}\omega_iE_i(\omega_j)+\epsilon_i(-\omega_i)) =$$

$$\mu(\epsilon_i(\omega_i)-\epsilon_i(-\omega_i) > -2h_i\omega_i-\sum_{j\neq i}2J_{i,j}\omega_iE_i(\omega_j)) \tag{16}$$

Using eq. (10), the logistic assumption for the errors, it is straightforward to manipulate this expression and conclude that

$$\mu(\omega_i \mid X_i, Y_{g(i)}, \mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i)) \propto exp(\beta_i h_i \omega_i + \sum_{j \neq i} \beta_i J_{i,j} \omega_i E_i(\omega_j)) \tag{17}$$

where " $\propto$ " means "is proportional to."

Moving from individual to joint conditional probabilities is now trivial, since the random utility terms are independent across individuals. The joint probability measure for the population choices is

$$\begin{aligned} &\mu(\underset{\sim}{\omega} \mid X_1, Y_{g(1)}, \ldots, X_I, Y_{g(I)}, \mu_1^e(\underset{\sim}{\omega}_{-1} \mid F_1), \ldots, \mu_I^e(\underset{\sim}{\omega}_{-I} \mid F_I)) = \\ &\prod_i \mu(\omega_i \mid X_i, Y_{g(i)}, \mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i)) \propto \\ &\prod_i exp(\beta_i h_i \omega_i + \sum_{j \neq i} \beta_i J_{i,j} \omega_i E_i(\omega_j)) \end{aligned} \tag{18}$$

Once one specifies the distributions across the population of private incentives, h_i, unobserved heterogeneity, β_i, interaction weights, $J_{i,j}$, and beliefs about the behaviors of others, $E_i(\omega_j)$, one has a complete characterization of the distribution of observed behaviors $\underset{\sim}{\omega}$. This explicit mapping of the distributions of individual characteristics and interdependence weights into a distribution of individual behaviors is the hallmark of interactions-based models.

In order to complete the description of the binary choice model, it is necessary to explore whether expectations can be rational in the sense described above. In the context of binary choice, rational expectations require that the beliefs $E_i(\omega_j)$ coincide with the mathematical expectations $E(\omega_j)$, $j = 1 \ldots I$ which are implied by probabilities described by eq. (18), conditioning on those beliefs. Recalling the definition of the hyperbolic tangent function, $tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$, one can express the expected value of each choice ω_i as

$$E(\omega_i \mid X_i, Y_{g(i)}, \mu_i^e(\underset{\sim}{\omega}_{-i} \mid F_i)) = tanh(\beta_i h_i + \sum_{j \neq i} \beta_i J_{i,j} E_i(\omega_j)). \tag{19}$$

A rational expectations equilibrium for this model requires that there exists a set of numbers $E(\omega_i)$ such that for all i and j

$$E(\omega_i) = tanh(\beta_i h_i + \sum_{j \neq i} \beta_i J_{i,j} E(\omega_j)) \tag{20}$$

Existence of a rational expectations equilibrium is thus a fixed point problem for the set of I equations described by (20). Fortunately, this is an easy case to analyze. Since the $tanh(\cdot)$ function is continuous with range $[-1,1]$, Brouwer's fixed point theorem for mappings may be immediately invoked to establish that at least one rational expectations solution exists.

ii. a baseline binary choice model

The basic binary choice interactions structure described in section 3.*i* can be specialized in many ways to incorporate different types of decisions, interactions environments and the like. In order to elucidate the general properties of models of this type, it is useful to consider a baseline case which has been extensively analyzed in Brock and Durlauf (2000a). This case assumes that for all i and j, 1) $h_i = h$, 2) $\beta_i = \beta$, and 3) $J_{i,j} = \frac{J}{I-1} \geq 0$.[10] Substantively, these assumptions do two things. First, the three assumptions eliminate all heterogeneity in individual behavior except that which is generated by the errors $\epsilon_i(\omega_i)$. Second, the social interactions have the property that each individual weighs the decisions of all others equally. This is obviously a very strong restriction on the nature of conformity effects.

Under these assumptions, the system of equations described by (20) reduces to

$$E(\omega_i) = tanh(\beta h + \frac{\beta J}{I-1} \sum_{j \neq i} E(\omega_j)) \ \forall \ i,j. \tag{21}$$

Since each agent is associated with the same parameters β, h, and J, one can show that all expectations $E(\omega_i)$ are equal. This in turn implies that the expected group average, $m = I^{-1} \sum_j E(\omega_j)$, equals $E(\omega_i)$ and that the expected average choices of others relative to i, $m_{-i} = (I-1)^{-1} \sum_{j \neq i} E(\omega_j)$ must equal this same number.

[10] Recall that I is the size of the population, so $I-1$ normalizes the interaction weights an agent applies to each of the other members of the group which is being modelled.

Therefore m will obey the functional relationship

$$m = tanh(\beta h + \beta J m) \tag{22}$$

Any m which is consistent with this equation is a possible equilibrium for expected average group-level behavior. At least one m exists which solves this equation, as discussed in section 3.*i*.

iii. multiple equilibria in the baseline model

Since the existence of an equilibrium has been established, the next question is to evaluate whether (22) possesses a unique solution. When multiple solutions exist, then the individual-level or micro-level structure of the model does not uniquely determine its macro-level characteristics. Why should one think that, in expectation, the average choice level is not uniquely determined? The answer lies in considering the implication of the assumption that $\frac{J}{I-1} \geq 0$. The magnitude of J influences the extent to which each individual makes a choice based on his beliefs concerning the choices of others. When this conformity effect is strong enough, it means that for many population members, the desire to conform to others dominates the other factors which influence choice. But when individual behavior is driven by a desire to be similar to others, this does not provide any information on what they actually do; rather it merely implies that whatever behaviors occur, there will be substantial within-group correlation due to conformity effects. Of course, the role of the conformity effect is determined by its strength relative to the private incentives agents face.

These considerations suggest that the number of equilibria should reflect an interplay of the various parameters of the model. Brock and Durlauf (2000a) contains the following theorem which characterizes the number of equilibria in this model.

Theorem 1: Relationship between individual behavioral parameters and number of self-consistent equilibria in the binary choice model with social interactions

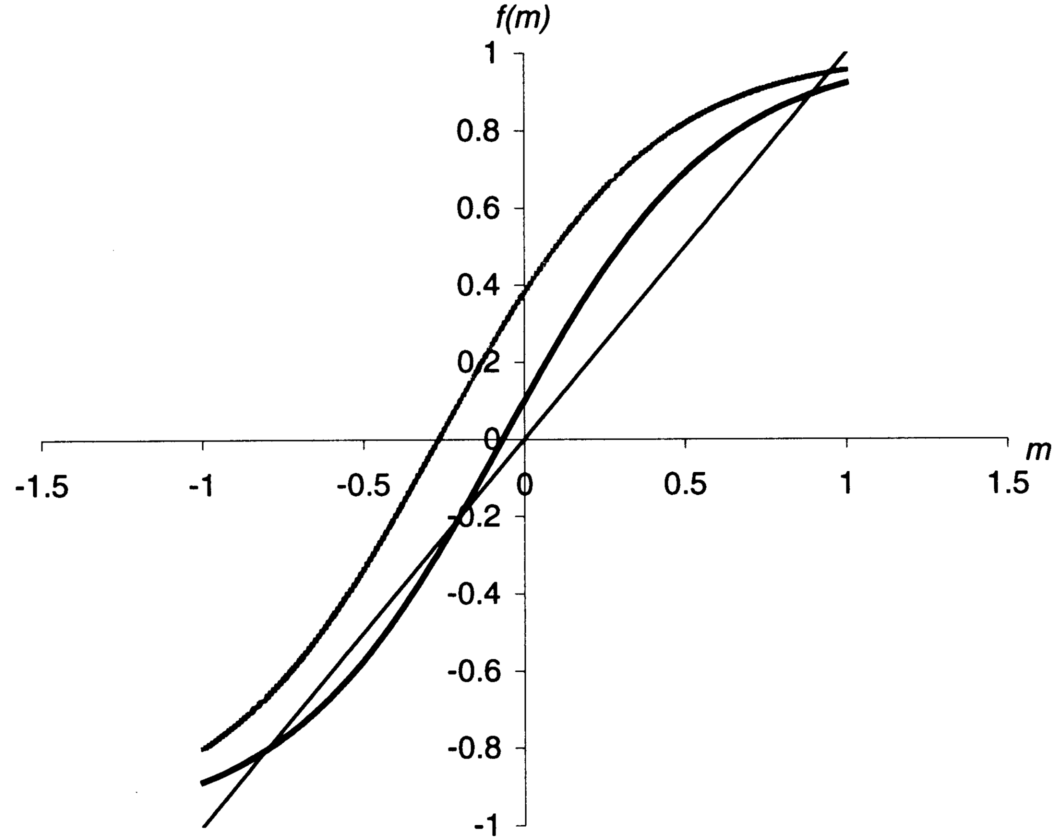

Figure 1. Equilibria for expected average choice

small h — large h

$$m = \tanh(\beta h + \beta Jm)$$
$$\beta J > 1$$

i. If $\beta J < 1$, then there exists a single solution to eq. (22).

ii. If $\beta J > 1$ and $h = 0$, there exist three solutions to eq. (22). One of these solutions is positive, one solution is zero, and one solution is negative.

iii. If $\beta J > 1$ and $h \neq 0$, there exists a threshold H, (which depends on βJ) such that

a. for $|\beta h| < H$, there exist three solutions to eq. (22), one of which has the same sign as h, and the others possessing opposite sign.

b. for $|\beta h| > H$, there exists a unique solution to eq. (22) with the same sign as h.

This theorem provides a description of the ways in which private incentives, h, unobserved heterogeneity, β, and social incentives, J, combine to determine the number of self-consistent equilibria in this population. Notice that it is the interplay of private and social incentives which controls multiplicity. Suppose that one fixes β and J so that $\beta J > 1$. In this case, different values of h will induce different numbers of equilibria. This is qualitatively illustrated in Figure 1 where equilibria of the model are defined by the intersection of the *tanh* with the 45° line.

A critical role is played by the βJ, in that large (in a sense specified in the theorem) values of this composite parameter are required for multiplicity. The role of J is easy to understand. Small values of J mean that the strength of endogenous interactions is week, which mitigates against self-consistent bunching of behavior. To understand the role of β, recall that small values of β imply that the likelihood of a large value of $|\epsilon_i(1) - \epsilon_i(-1)|$ is relatively high. Hence for small β's, a relatively large percentage of the population will have draws of $\epsilon_i(1) - \epsilon_i(-1)$ which, roughly speaking, dominate the deterministic parts of their payoffs. Further, the expected value of the percentage who are led to choose 1 due to the random utility draws will, in expectation, be the same as the expected value of the percentage of the population led to choose -1. In other words, small β's imply that a relatively small percentage of the population is susceptible to self-consistent bunching, in the sense that their decisions will, marginally be decided by the social utility component of their payoff functions. Now, consider what is needed for multiple self-consistent equilibria. Intuitively, what is needed is that enough of the population is susceptible to being influenced by the expected choices of others. But this requires, for fixed J, relatively large values of m. When the percentage of individuals whose behavior is dominated by the unobserved heterogeneity is high enough, then the range of possible values of m is restricted, which precludes the self-consistent bunching at multiple levels.

Under standard dynamic analogs to this static model, it turns out that the extremal equilibria are locally stable, whereas the interior equilibrium (measured in terms of the value of m) is not.[11] One can ignore this interior equilibrium as it cannot be expected to arise in practice. Thus, when multiple equilibria are present, we can restrict attention to the possible equilibrium m^*_+ in which the average choice is expected to be positive and m^*_- in which the average choice is expected to be negative.

How is a particular equilibrium selected? The model, as specified, does not provide an answer to this question. Within game theory, there is a vast literature on the general question of equilibrium selection in models with multiple equilibria; see Samuelson (1997) for an overview. The analysis of equilibrium selection typically requires formulating a Markov version of the sort of model I have described. Unlike the analysis of stability described above, noise in the individual decision processes is introduced to allow the population to shift across equilibria, so the relevant question becomes the percentage of time a given equilibrium is realized over the sample path of the process. Blume and Durlauf (2000) do this for the Brock-Durlauf (2000a) model and find that over very long horizons, the set of aggregate choices will tend to cluster around the welfare-superior equilibrium. This result does not say that the welfare-inferior equilibrium is uninteresting, as the mean first passage time from that equilibrium to the welfare-superior one can be extremely long. What this means is that in the presence of social interactions, initial conditions may have particularly persistent effects, i.e. history matters.

***iv.* social welfare in the baseline model**

[11]These dynamic analogs typically assume that agents at t react to the expected average choice for period $t-1$.

The presence of multiple equilibria naturally leads to the question of the relationship between a particular equilibrium and aggregate social welfare. In the baseline case of common h, β, and J values, it is natural to ask which equilibria produces the highest expected average payoff for the population. Brock and Durlauf (2000a) prove the following theorem.

Theorem 2. Welfare rankings of equilibria

i. When $h > 0$ (< 0), then the equilibrium associated with m^*_+ (m^*_-) provides a higher level of expected utility for each agent than the equilibrium associated with m^*_- (m^*_+).

ii. When $h = 0$, then the equilibrium associated with m^*_+ and the equilibrium associated m^*_- provide equal levels of expected utility for each agent.

In words, when $h \neq 0$, then individuals are better off when the average choice is of the same sign as h. Intuitively, when this holds, the social incentives to conform work in the same direction as the private incentives. On the other hand, when $h > 0$ and $m = m^*_-$, then these incentives clash. In this case, the average member of the population would be better off if the other equilibrium prevailed. The existence of a welfare inferior equilibrium illustrates how individually rational decisions can be collectively undesirable.

4. Conceptual issues

i. methodological individualism

Interactions based models represent an effort to introduce richer sociological structure into economic theory. The absence of such structures has been the source of severe criticism of economic theory by social scientists who are not economists as well as among heterodox economists. Granovetter (1985) gives a typical critique:

"Classical and neoclassical economics operates, in contrast, with an atomized and *under*socialized conception of human action...The theoretical arguments disallow by hypothesis any impact of social structure and social relations on production, distribution, or consumption." (pg. 55)

The interactions-based approach represents one way of answering this criticism without sacrificing any of the basic microeconomic behavioral assumptions of economics. I make this claim in two respects: one superficial and the other somewhat deeper.

First, the modelling exercise described in Sections 2 and 3 illustrates how one can formally integrate group-level influences into individual decisions so long as these influences can be modelled as variables, as is done by assuming that individual payoffs depend on $Y_{g(i)}$ and m, for example. Self-consistency conditions of the type we have modelled impose those feedbacks that exist between the choices of the members of a group, in a way analogous to the relationship between the modelling of individual demand schedules as functions of prices and the requirement that these prices clear markets.

Of course, the mathematics assumes that social components of behavior can be translated into the variables and behavioral parameters I have employed; I interpret part of Abbott's (1988,1997) criticism of "variables-based" models as essentially asserting that reduction cannot occur. In my opinion, one cannot make a compelling critique of this so-called variables approach outside of specific research contexts. The judgment as to whether or not a particular mathematical model provides understanding of a given phenomenon depends on one's background knowledge of the phenomena (which will include qualitative factors such as ethnographic studies) as well as one's knowledge of the mathematical properties of the model under study (what properties of a model are robust to various alternative modelling assumptions, etc.) Put differently, the evaluation of social science research cannot be reduced to an algorithm; false assumptions, oversimplifications and the like can only be evaluated in the context of an overall research design.[12] Nevertheless, the exercise illustrates that there is no reason in principle that methodological individualism cannot incorporate

[12] It is for this same reason that many of Freedman's (1987,1991) criticisms of causal modelling in the social sciences seem to be excessive; contrast these with Sobel (1998) whose criticisms are appropriately contextualized.

the socialized forms of behavior to which Granovetter refers.

This perspective on methodological individualism is quite similar to that taken by Elster (1989) in the context of studying social norms:

"...I believe one can define, discuss, and defend a theory of social norms in a wholly individualistic framework. *A norm, in this perspective, is the propensity to feel shame and to anticipate sanctions by others at the thought of behaving in a certain, forbidden way*...this propensity becomes a *social* norm when and to the extent that it is shared with other people...the social character of the norm is also manifest in the existence of higher-order norms that enjoin us to punish violators of the first-order norm. To repeat, this conception of a network of shared beliefs and common emotional reactions does not commit us to thinking of norms as supraindividual entities that somehow exist independently of their supports." (pg. 105-106)

Second, the modelling framework I have described illustrates how social interactions lead to features of group behavior which are qualitatively different from those which arise in environments without such interactions. This is an example of the property of "emergence" which is a common feature of environments of this type. Emergence leads to the question of the relationship between these models and statistical mechanics models in physics, which is discussed next.

ii. statistical mechanics and social science

The model of binary choice with social interactions which I have described lies in a class of mathematical models which originate in physics, specifically in the area of statistical mechanics.[13] These models were originally developed to explain how magnets arise in nature. A magnet occurs when, for a piece of iron, a majority of the atoms spin up or spin down (spin is a binary property of atoms.) In the early twentieth century, the existence of natural magnets was a major puzzle, as there are physical reasons why the *ex ante* probability that a given atom spins up or down is .5. The resolution of this puzzle, first instantiated in the celebrated Ising-Lenz model, is to assume that the probability that a given atom possesses a certain spin depends on the spins of the nearest neighbors to that atom. The subsequent statistical mechanics literature has extended analyses of these types to a wide range of alternative interaction structures which correspond to different choices of $J_{i,j}$.

[13] Yeomans (1982) is an accessible introduction to statistical mechanics.

For the purposes of social science, of course, the physical interpretation of these models is of no interest. What is of enormous interest are the mathematical properties of these systems, which have made the same mathematical models valuable in areas ranging from computer science (where neural networks have this mathematical structure) to biology (models of molecular evolution); Anderson and Stein (1984) is an accessible discussion of this. Among the many interesting properties of these systems are: emergence, symmetry breaking, nonergodicity, phase transition, and universality. I provide a very brief description of each property to clarify how it arises in the model that has been developed.

a. emergence[14]

In a system of interacting agents, emergent properties are those which cannot be reduced to statements about the individual elements when studied in isolation. In physics, magnetism is an emergent phenomenon as it is a property of many iron atoms whose atomic spins are aligned; similarly, ice is a property of the way in which many water molecules are arrayed, not one molecule in isolation. The multiple equilibria which are described in Theorem 1 are examples of emergence in a socioeconomic context, as they constitute a property which arises only with respect to a group rather than for a single individual. Another example of emergence is Schelling's (1971) celebrated demonstration of how complete segregation is produced from mildly discriminatory preferences.

One important aspect of emergence is that it breaks any logical relationship between methodological individualism and reductionism. What I mean is that emergent properties cannot be understood through the individual elements of a system, even though the behaviors of these elements matter for the emergent properties.

b. symmetry breaking

Symmetry breaking occurs when for a specification of symmetrically specified agents, asymmetric outcomes occur. To see how this model exhibits symmetry-

[14]See Anderson (1982) for a physical science perspective on emergence.

breaking, consider the case $h = 0$. In this case, each agent is *ex ante* privately indifferent between the two choices (i.e. in the absence of any social interaction effect, the probability that an agent chooses 1 (and of course by implication -1) is .5. However, when $J > 1$, the choices will bunch (in expected value) around one of the choices. Symmetry breaking is important in modelling spatial agglomeration of agents (see Fujita for a neoclassical approach and Krugman (1996) for a statistical mechanics perspective) into regions and cities.

c. nonergodicity

A probabilistic system is nonergodic if the conditional probabilities which describe the behavior of each element of the system conditional on the other elements fail to uniquely characterize the behavior of the system as a whole. The cases where (22) exhibits multiple solutions are thus nonergodic. In these cases, one specifies the conditional probability choices of each member of the population and imposes self-consistency; however, this does not uniquely determine the aggregate behavior of the population. Multiple equilibria are a common feature of coordination games in economics, which are noncooperative environments in which individuals conform to one another – see Cooper (1997) for an overview. Such models typically embody conformity effects of the type which have been described.

***d*. phase transition**

A model exhibits a phase transition when a small change in a model parameter induces a qualitative shift in the model's properties. The binary choice with social interactions model exhibits phase transitions along two dimensions. Recalling Theorem 1, holding $\beta J > 1$ constant, there is a threshold value H (which depends on βJ) such that if $|\beta h|$ moves from less than H to greater than H, the number of equilibria shifts from 3 to 1. On the other hand, the theorem also implies that if one holds $|\beta h|$ constant, then there exists a threshold K (which depends on $|\beta h|$) such that as βJ moves from less than K to greater than K, the number of equilibria shifts from 1 to 3. These shifts in the number of equilibria are classic examples of phase transitions.

e. universality

A universal property of a system is one which does not depend on details of the system's micro-level specification. Such properties are found in many physical contexts; for example, magnetization of the type captured in the Ising-Lenz model does not depend on the nearest neighbor interaction structure. Universality is extremely appealing from the perspective of social science modelling, since we often do not have any real justification for choosing interactions structures, forms of interaction effects, etc. outside of analytical convenience.

To see how this model exhibits some types of universality, suppose that each individual is associated with a neighborhood n_i which characterizes the set of individuals in the population with whom he wishes to conform; $\#(n_i)$ denotes the neighborhood's population size. Assume as well that each member of the neighborhood receives an equal conformity weight, so that $J_{i,j} = J/(\#(n_i) - 1)$. A self-consistent equilibrium for this system is any set of solutions $E(\omega_1)...E(\omega_I)$ to the set of I equations

$$E(\omega_i) = tanh(\beta h + \frac{\beta J}{\#(n_i) - 1} \sum_{j \in n_i} E(\omega_j)) \tag{23}$$

This mapping must possess at least one fixed point and hence at least one self-consistent equilibrium exists. Notice that any solution of the model with global interactions must also represent a solution to this model; specifically, it is always possible that all expectations are identical, so that

$$E(\omega_1) = E(\omega_2) = E(\omega_I) = m = tanh(\beta h + \beta J m). \tag{24}$$

Hence the properties we have found for the global interactions model occur for a wide variety of alternative interaction structures. This being said, universality has been relatively unexplored in economics applications of statistical mechanics. My judgment is that this is an important area for future work.

iii. social capital

This model of social interactions and individual choice also has the potential to facilitate a more rigorous analysis of social capital. As powerfully argued by Portes (1998) among others, a major problem with the social capital literature is the lack of definitional clarity. Putnam's recent exhaustive empirical account (2000), for example, relies on functional definitions of what social capital is. As I have described in Durlauf (1999), my reading of the social capital literature suggests that an operational definition of social capital is "the influence which the characteristics and behaviors of one's reference groups has on an individual's assessments of alternative courses of behavior." If this definition is reasonable, then social capital, at least to a first approximation, can be measured by the $J_{i,j}$ terms in the individual choice functions.

5. Statistical implementation

In bringing social interactions models to data, a number of difficult statistical issues arise. In particular, under the assumption of rational expectations, there exist relationships between the various regressors which comprise the model. These relationships, in turn, influence identification. This possibility was first recognized by Wallis (1980) in the context of time series models. A seminal paper by Manski (1993) has developed this idea in the context of interactions-based models.

i. linear-in-mmeans model

To see how an identification problem arises in interactions-based models, it is useful to consider a linear regression model that is analogous to the theoretical model I have described. This model, known as the linear-in-means model, is the primary one used in empirical work on interactions and so is of independent interest as well. For this model, which is a generalization of the one studied by Manski (1993)[15], each agent is, as before assumed to be a member of a group $g(i)$ whose contextual and endogenous characteristics both affect the individual. However, individuals in the data sample are not required to be members of a common group. This means that each individual in the data set may be associated with a distinct expected choice level $m_{g(i)}$, where this expectation is made conditional on some information set $F_{g(i)}$. Choices are continuous and are represented by a regression equation,[16]

$$\omega_i = k + c'X_i + d'Y_{g(i)} + Jm_{g(i)} + \epsilon_i. \tag{25}$$

To see how an identification problem arises for the linear-in-means model, it is useful to work with the reduced form of the model, which means replacing the unobservable term $m_{g(i)}$ with its value as determined by the model. If one takes the expected value of each side of (25) with respect to $F_{g(i)}$, and rearranges terms, $m_{g(i)}$ can be shown to follow

$$m_{g(i)} = \frac{k + c'E(X_i \mid F_{g(i)}) + d'E(Y_{g(i)} \mid F_{g(i)})}{1-J} \tag{26}$$

which can be substituted into (25) to produce

$$\omega_i = \frac{k}{1-J} + c'X_i + \frac{1}{1-J}d'E(Y_{g(i)} \mid F_{g(i)}) + \frac{J}{1-J}c'E(X_i \mid F_{g(i)}) + \epsilon_i. \tag{27}$$

This regression is estimable by ordinary least squares. A potential identification problem occurs because of the possible linear dependence between $E(Y_{g(i)} \mid F_{g(i)})$ and $E(X_i \mid F_{g(i)})$. For example, in the model studied by Manski (1993), it is assumed that $E(Y_{g(i)} \mid F_{g(i)}) = E(X_i \mid F_{g(i)})$, i.e. that the contextual effects which affect individuals are equal to expectations of the neighborhood averages of the same variables which affect them on an individual level. This is the source of the nonidentification result found by Manski (1993).

On the other hand, this formulation makes clear that identification is possible

[15]In Manski (1993), $Y_{g(i)}$ is assumed to equal $X_{g(i)}$, the average of X_i across all members of $g(i)$.

[16]The linear-in-means model is typically not developed from an explicit choice-based framework of the type described in Section 2, unlike the binary choice model.

in the linear-in-means model. In the presence of prior information on which individual and contextual variables influence individual behavior, it is possible that $E(Y_{g(i)} \mid F_{g(i)})$ and $E(X_i \mid F_{g(i)})$ are linearly independent, which will mean the model is identified. For example, suppose that $F_{g(i)}$ includes $Y_{g(i)}$ and $X_{g(i)}$, where the latter denotes the neighborhood $g(i)$ average of X_i. The reduced form for the individual behavioral equation is in this case

$$\omega_i = \frac{k}{1-J} + c'X_i + \frac{1}{1-J}d'Y_{g(i)} + \frac{J}{1-J}c'X_{g(i)} + \epsilon_i. \tag{28}$$

Assuming that X_i contains r elements and $Y_{g(i)}$ contains s elements, this regression has $2r+s+1$ regressors and $r+s+2$ unknowns; it is easy to verify that if the regressors are all linearly independent then the system is identified if $r>0$ and overidentified if $r>1$. Further, suppose that $X_{g(i)}$ is expressed as

$$X_{g(i)} = \Pi_0 + \Pi_1 X_i + \Pi_2 Y_{g(i)} + \eta_i \tag{29}$$

This allows us to rewrite the individual reduced form as

$$\omega_i =$$

$$\frac{k}{1-J} + \frac{Jc'}{1-J}\Pi_0 + (c' + \frac{Jc'}{1-J}\Pi_1)X_i + (\frac{J}{1-J}d' + \frac{Jc'}{1-J}\Pi_2)Y_{g(i)} + \frac{J}{1-J}c'\eta_i + \epsilon_i. \tag{30}$$

Equation (29) identifies Π_0, Π_1, and Π_2. Since η_i is orthogonal to the other regressors in (30), coefficients which equal $\frac{k}{1-J} + \frac{Jc'}{1-J}$, $c' + \frac{Jc'}{1-J}\Pi_1$, and $\frac{J}{1-J}d' + \frac{Jc'}{1-J}\Pi_2$ respectively are identified from a regression of ω_i onto a constant, X_i, and $Y_{g(i)}$. Hence, the coefficients in (30) define a set of $r+s+1$ equations with $r+s+2$ unknowns when η_i is omitted. In order to identify the structural coefficients, it is necessary that the regressors η_i provide an additional estimate of some component of the vector $\frac{J}{1-J}c'$. This will give as many coefficients as there are unknowns. This, in turn, requires that η_i is not identically equal to zero, i.e. that there is some part of the neighborhood averages of the individual controls which do not lie in the space spanned by a constant, X_i, and $Y_{g(i)}$. In turn, a necessary condition for η_i to be nonzero is that the structural model of individual decisions, eq. (25), contains at least one regressor x_j such that its neighborhood average $\overline{x}_{j,g(i)}$ is excluded from $Y_{g(i)}$ prior to the exercise.

***ii.* nonlinear models**

Notice that linearity plays a critical role in creating the potential for nonidentification. Suppose that instead of (25), the individual-level behavioral equation is

$$\omega_i = k + c'X_i + d'Y_{g(i)} + J\phi(m_{g(i)}) + \epsilon_i. \tag{31}$$

for some invertible function $\phi(\cdot)$. The rational expectations condition for this model is:

$$m_{g(i)} = \psi(k + c'E(X_i \mid F_i) + d'E(Y_{g(i)} \mid F_i)). \tag{32}$$

where $\psi(\cdot) = (1-\phi(\cdot))^{-1}$. The associated reduced form equation equals

$$\omega_i = k + c'X_i + d'Y_{g(i)} + J\phi(\psi(k + c'E(X_i \mid F_i) + d'E(Y_{g(i)} \mid F_i))) + \epsilon_i. \tag{33}$$

This is an example of a partially linear model (cf. Horowitz (1998)); Brock and Durlauf (2000b) verify that the parameters of this model are locally identified under weak assumptions on the variables in the system.

***iii.* binary choice**

Using the linear model as a baseline, I now consider identification in the binary choice model. In order to do this, I focus on the analogous binary choice model with social interactions. This model will possess a likelihood function

$$L(\underset{\sim}{\omega}_I \mid \underset{\sim}{X}_i, \underset{\sim}{Y}_{g(i)}, m^e_{g(i)} \ \forall\ i) =$$

$$\prod_i \mu(\omega_i = 1 \mid \underset{\sim}{X}_i, \underset{\sim}{Y}_{g(i)}, m^e_{g(i)})^{\frac{1+\omega_i}{2}} \cdot \mu(\omega_i = -1 \mid \underset{\sim}{X}_i, \underset{\sim}{Y}_{g(i)}, m^e_{g(i)})^{\frac{1-\omega_i}{2}} \sim$$
$$\prod_i (exp(\beta k + \beta c' \underset{\sim}{X}_i + \beta d' \underset{\sim}{Y}_{g(i)} + \beta J m^e_{g(i)})^{\frac{1+\omega_i}{2}}$$
$$\cdot\, exp(-\beta k - \beta c' \underset{\sim}{X}_i - \beta d' \underset{\sim}{Y}_{g(i)} - \beta J m^e_{g(i)})^{\frac{1-\omega_i}{2}}) \tag{34}$$

where rational expectations imposes the restriction

$$m^e_{g(i)} = m_{g(i)} = \int tanh(\beta k + \beta c' X + \beta d' Y_{g(i)} + \beta J m_{g(i)}) dF_{X \mid Y_{g(i)}}. \tag{35}$$

under the assumption that agents only know $Y_{g(i)}$ within a neighborhood.[17] Notice the multiplicative structure of the parameters of the model; this makes it is necessary to normalize the parameters in order to achieve identification, this can be done by setting $\beta = 1$.

The key question in terms of identification of the linear-in-means model is whether $m_{g(i)}$ is collinear with the other regressors in the individual behavioral equation due to self consistency. This same issue arises in the binary choice case. However, in the binary choice model, the expected value of a neighborhood choice is a nonlinear function of the other variables in the model. The Mathematical Appendix gives a formal statement of the conditions under which the binary choice model with social interactions is identified, but the key intuition for identification follows from the nonlinearity built into (35). As the theorem indicates, there is no need for an exclusion restriction on the contextual variables in order to achieve identification, as was true in the linear-in-means model. This is an example of a more general phenomenon noted by McManus (1992), namely that lack of identification in parametric systems is typically associated with linearity.

This being said, the standard errors in estimating the binary choice model may

[17] It is technically convenient, in working with the binary choice model, to assume that the information sets for individuals take this form, rather than to assume each individual knows the distribution of X_i within his group, as was done in the linear case. This is so because, as seen in Manski (1988), identification arguments in the binary choice model are more subtle than for linear regressions. See Brock and Durlauf (2000b) for details.

be extremely large if (35) is not "too" nonlinear. Exclusion restrictions of the type which generate identification in the linear model can facilitate accurate estimation for this case.

iv. self-selection

The discussion of identification has thus far treated the neighborhood variables as uncorrelated with the error terms in the equations describing individual behavior. For contexts such as residential neighborhoods, this assumption is problematic. However, it turns out that paradoxically, self-selection, when properly controlled for, can assist in achieving identification.

In this example, suppose one had a data set of Catholic schools and that one wanted to assess the effects of characteristics of the schools on educational attainment. Assume these schools provide both peer group and contextual influences, so Catholic schools are an example of the types of neighborhoods we have been discussing, i.e. that eq. (25) describes individual outcomes.

In order to analyze how the endogeneity of the decision to use Catholic schools affects inference concerning interaction effects, I follow the framework developed by Heckman (1979) in a classic article; my development is taken from Brock and Durlauf (2000b). Assume that the family of each child in the data set has made a binary choice whether or not to send the child to Catholic school based on a latent variable z_i which measures a family's evaluation of the relative merits of the school. A child is a member of the data set if and only if $z_i > 0$. In turn, this latent variable is assumed to have the form

$$z_i = \gamma' R_i + \eta_i \tag{36}$$

where R_i is a vector of determinants of i's evaluation of the school. Finally, assume that the errors ϵ_i and η_i are jointly normal with conditional expected values $E(\epsilon_i \mid X_i, Y_{g(i)}, m_{g(i)}, R_i) = E(\eta_i \mid X_i, Y_{g(i)}, m_{g(i)}, R_i) = 0$ and variance/covariance matrix

$$\begin{bmatrix} \sigma_\epsilon^2 & \rho\sigma_\epsilon \\ \rho\sigma_\epsilon & 1 \end{bmatrix}. \tag{37}$$

These assumptions are sufficient to imply that

$$E(\epsilon_i \mid z_i > 0) = \rho\sigma_\epsilon\lambda(\gamma' R_i) \tag{38}$$

where $\lambda(\gamma' R_i)$ is defined as

$$\lambda(\gamma' R_i) = \frac{\phi(\gamma' R_i)}{\Phi(\gamma' R_i)}, \tag{39}$$

with $\phi(\cdot)$ and $\Phi(\cdot)$ denoting the standard normal density and distribution functions respectively. A regression whose residual is orthogonal to the various regressors can therefore be constructed by adding $\lambda(\gamma' R_i)$ as a regressor, since for the linear equation

$$\omega_i = k + c' X_i + d' Y_{g(i)} + J m_{g(i)} + \rho\sigma_\epsilon\lambda(\gamma' R_i) + \zeta_i. \tag{40}$$

ζ_i is orthogonal to the other variables on the right hand side of (40). The fact that a selection correction can be formulated as an additional regressor is the critical insight of Heckman (1979). For our purposes, the key question is whether the parameters of this equation are identified.

To see how the correction affects identification, it makes sense to consider the case where each individual control is matched one-to-one with a contextual effect so that $X_{g(i)} = Y_{g(i)}$. Assume as well that none of the variables in R_i are functionally dependent on $m_{g(i)}$, so that we may assume that the reduced form for $m_{g(i)}$ depends on R_i. As discussed above, if $\rho\sigma_\epsilon = 0$, so there is no selection correction, the regression then corresponds to Manski's (1993) nonidentification example. However, in the presence of self-selection, the expected average choice within a neighborhood is, under self-consistency

$$m_{g(i)} = \frac{k}{1-J} + \frac{1}{1-J}(c' + d') Y_{g(i)} + \frac{\rho\sigma_\epsilon}{1-J} E(\lambda(\gamma' R_j) \mid j \in g(i)) \tag{41}$$

so that a reduced form for individual behavior may be written as

$$\omega_i =$$

$$\frac{k}{1-J} + c'X_i + \frac{1}{1-J}(Jc' + d')Y_{g(i)} + \rho\sigma_\epsilon\lambda(\gamma' R_i) + \frac{J\rho\sigma_\epsilon}{1-J}E(\lambda(\gamma' R_j) \mid j \in g(i)) + \zeta_i \quad (42)$$

In this reduced form regression, nonidentification when $\rho\sigma_\epsilon = 0$ follows immediately from observing that there are $2r+2$ parameters and only $2r+1$ regressors. However, when there is a selection correction, two new regressors are introduced, $\lambda(\gamma' R_i)$ and $E(\lambda(\gamma' R_j) \mid j \in g(i))$, but only one new parameter, $\rho\sigma_\epsilon$. This allows for identification so long as $\lambda(\gamma' R_i)$ and $E(\lambda(\gamma' R_j) \mid j \in g(i))$ are not perfectly collinear, which requires that there is within-school variation in $\lambda(\gamma' R_i)$. Notice that the nonlinearity of $\lambda(\,\cdot\,)$ ensures that the appearance of regressors in R_i which appear in either X_i or $Y_{g(i)}$ does not imply nonidentification due to multicollinearity of the correction term with the other variables in the model.

Recall from our earlier discussion that in order to achieve identification in the linear model, one needs an individual control whose school average is not a contextual effect. This is precisely what occurs when $\lambda(\gamma' R_i)$ is introduced into the linear-in-means model, since $E(\lambda(\gamma' R_j) \mid j \in g(i))$ is not an element of the model even when selection is controlled for.

When $m_{g(i)}$ is a component of R_i, an alternative route to identification may be developed. For ease of exposition, suppose that the expected average choice level is the only element in R_i. The selection-corrected linear-in-means model is now

$$\omega_i = k + c'X_i + d'Y_{g(i)} + Jm_{g(i)} + \rho\sigma_\epsilon\lambda(\gamma m_{g(i)}) + \zeta_i. \quad (43)$$

The parameters in this regression will now be identified so long as the joint support of X_i and $Y_{g(i)}$ does not lie in a proper linear subspace of R^{r+s} since the nonlinearity of the selection correction ensures that there is no linear dependence between $m_{g(i)}$ and the individual and neighborhood controls. Notice that this is the same reason for identification derived for the binary choice model; in both cases, the nonlinear dependence of $m_{g(i)}$ on X_i and $Y_{g(i)}$ produces identification.

Of course, identifiability of model parameters does not say anything about the precision of the estimates facilitated by selection corrections. Intuitively, one will need substantial cross-school variation in $m_{g(i)}$ if the nonlinear dependence of the correction on this term is the basis for identification. Similarly, substantial variation in R_i may be needed if elements of this vector are highly correlated with combinations of X_i and $Y_{g(i)}$, in order for the nonlinear selection correction to avoid near-multicollinearity within a given sample. Notice in this case, the presence of regressors in R_i which do not appear in X_i or $Y_{g(i)}$ will likely prove valuable in practice.

6. Evidence

While structural models of social interactions are in principle estimable, they have yet to be implemented empirically. Empirical support for models of this type must therefore be sought in other contexts.

i. social psychology

The most compelling evidence of the role of social interactions in individual behavior has been produced by social psychologists. While many demonstrations of the salience of social interactions may be found in the social psychology literature, perhaps the most impressive study is the celebrated Robbers Cave experiment, described in Brown (1986) as "the most successful experiment ever conducted on intergroup conflict." This experiment is described in Sherif et al (1961). Sherif and collaborators studied the behavior of a group of teenage boys at an isolated retreat in Robbers Cave State Park in Oklahoma. A group of boys were initially placed in a common living quarters and associated social environment. Once friendships and other social relations developed, the experimenters announced that the boys were assigned to two groups, Rattlers and Eagles. The new assignments were essentially random, with the exception that strong friendship pairs were broken up. A set of competitive activities were initiated. Sherif et al (1961) documents in great detail how the two groups developed strong internal senses of identity along with great animosity towards the other group, animosity which carried over beyond the competitive activities. Previous friendships disappeared, attribution of negative stereotypes to the other group became commonplace. While the introduction of cooperative activities diminished the hostility, the experiment clearly demonstrated that group identification can strongly influence individual behavior.

ii. historical evidence

A second type of evidence comes from many historical studies; and one need not go back to Ancient Greece to identify the interplay of private and social factors in social, economic, and political outcomes (although evidence from the classical period can certainly be instructive). Browning (1992) provides an in-depth study of the behavior of members of a German order battalion during World War II. This battalion consisted of older Germans who were used as reserve troops and eventually were ordered to kill civilians, primarily Jews. Interestingly, it turns out that not only were there no incidents of formal punishment for refusing such orders, but it was known that this was so. What Browning found was that about 20% of the 300 men he studied refused to obey genocidal orders, despite the pressures brought on them to do so. What seems to have differentiated this 20% from the others is not differences in politics or education, but rather the fact that they had Jewish friends or social contacts with Jews prior to the war. Hence prior social interactions had dramatic effects on behavior. Put differently, in understanding the variety of behaviors of the soldiers in Browning's study, one needs to account for different contextual experiences which determined how they weighed the pressures to act which were common to all of them.

Obviously, one can object that this example has little directly to do with social pathologies such as inner city violence. After all, the willingness of civilians to commit atrocities in World War II explores human nature in a far more extreme environment than any American inner city. However, Browning's work is useful in my judgment as it represents the equivalent of a split brain study in neuroscience. Neuroscience has made many advances based upon studying severely damaged brains (an example is when the two hemispheres are split); such damage implicitly gives insights into how normal brains function. My belief is that studies such as Browning's are very suggestive of how social environments condition behavior in general. By implication, when one is offering policy advice for particular contexts, this sort of information would seem to be relevant for the sort of priors one uses in the exercise.

More generally, it is difficult to interpret any number of major historical events such as 1) the resilience of Athenian democracy institutions and norms in the face of plague, massive war casualties and military occupation during and in the aftermath of the Peloponnesian War (Finley (1985), Ober (1989)), 2) the sudden outbreak of nearly fifty distinct democratic and workers revolutions in Europe in 1848 (Robertson (1952), Sperber (1994)) and 3) the sudden disintegration of the Czarist authority temporarily in 1905 (Ascher (1988,1992)) and permanently in 1917 (Figes (1997)), without explicit attention to the interrelationships of private incentives and social influences. It is therefore difficult to justify ignoring such an interplay in socioeconomic contexts such as the formation and perpetuation of slums, the breakdown of two parent families, drug abuse, and a host of social pathologies which are the primary focus of recent poverty research.

iii. interventions in group memberships – natural experiments

A third source of evidentiary support for interactions-based models, and one that is much closer to the substantive phenomena for which interactions-based models have been developed, comes from studies of the consequences of government interventions on various group memberships. These types of studies are now quite popular in economics, and often are referred to as "natural experiments." The source of this terminology is that ideally (from the perspective of inference) such interventions can be modelled as randomized experiments with respect to who receives which treatment, where treatments in this case mean exposure to different group influences.[18]

The most discussed intervention of this type is the Gautreaux demonstration. The basic history of the Gautreaux program is described in Rubinowitz and

[18]Heckman (1997,2000) provides a good discussion of the rationale for studying such data sets as well as an incisive analysis of the limits of the analogy between data of this type and actual randomized experiments one finds in medicine, etc.

Rosenbaum (2000). The Gautreaux program came into existence as a result of a successful law suit by public housing residents against the Chicago Housing Authority, who alleged that the location of their public housing was discriminatory because of their consignment to high crime areas. The program in essence offered relief to these plaintiffs by creating opportunities for families to move, depending on when a vacancy occurred, to either another housing unit in Chicago or to an adjacent suburb. Families who were eligible for relief were randomly contacted when units became available. As families had no control over whether they were offered a chance to move in or out of Chicago, and since essentially all families who were offered a move accepted, it is reasonable to treat the assignment of families to suburbs as a random treatment.

Sociologist James Rosenbaum, in conjunction with several coauthors, has tracked the outcomes of a number of the families who were moved under the program. As documented in Rosenbaum (1995), Rosenbaum, DeLuca and Miller (1999) and Rosenbaum and Popkin (1991) and elsewhere, striking differences exist between the outcomes for families who stayed in Chicago versus those who moved to suburbs. For example, children from families who were moved to suburbs had substantially higher high school graduation rates and post graduation wages than their counterparts who moved within Chicago.

The Gautreaux evidence is, in my judgment, important evidence of neighborhood effects, but two caveats should be noted in interpreting it. Each concerns self-selection. First, there was some self-selection in program participation. In order to be eligible for a move, a family had to have a history of timely payments on their previous public housing and to have maintained the quality of the housing. This eliminated 30% of those families who were eligible. One must therefore be careful about extrapolating the effects of this program to populations of poor people who have not been so screened. Second, the Rosenbaum data do not include records of individuals who moved to suburbs and subsequently returned. This means the data represent a self-selected sample of families. To see why this makes interpretation of the results problematic, suppose, for example, that all families found the loss of proximity to family and friends equally onerous, but that families differ in the weight attached to offspring educational achievement. In this case, the Gautreaux data would suffer from the problem that the neighborhood effects of the suburbs are entangled with the self-selection of relatively "good" parents. To be clear, there is no obvious way to measure the impact of the bias associated with these caveats.[19]

A second major quasi-experiment is the Moving to Opportunity Demonstration, currently being conducted by the Department of Housing and Urban Development to evaluate the effects of moving low-income families out of high-poverty neighborhoods; a detailed discussion of the program appears in Goering (1996). The demonstration randomly assigned a set of low income families who applied to participate in the program to one of three groups: 1) a group whose members are eligible for housing vouchers which provide a rent subsidy but which are only usable in census tracts with less than 10% poverty, 2) a group that is eligible for housing vouchers with no locational restrictions, and 3) a group in which members are not eligible for housing vouchers but can still receive public housing assistance based on previous eligibility. The demonstration is being conducted in 5 metropolitan areas: Baltimore, Boston, Chicago, Los Angeles, and New York City.

Preliminary results on the various experiments are becoming available. Ladd and Ludwig (1998) report evidence that those families in Baltimore that moved out of low income census tracts achieved better access to better schools as measured by a range of criteria. However, they find little evidence that the value added of these schools for the children in these families is higher than the schools used by families in the comparison and control groups. For the Boston demonstration, Katz, Kling, and Liebman (1999) similarly find evidence that the MTO program has been successful in generating relocation of families to better communities. They also find that children in both classes of families which were eligible for vouchers exhibited higher test scores as well as lower incidences of health and behavioral problems.

Like Gautreaux, the MTO demonstration suffers from a range of problems which make interpretation of the data difficult. First, the program is far from a pure randomized experiment, i.e. an experiment where living in a low poverty neighborhood constitutes a treatment. As mentioned before, participation in the program required families to volunteer; further, many families who were eligible to move to more

[19] It would be interesting to see if a bound on the effects of suburbs on Gautreaux families could be constructed using methods developed by Manski (1995). Brock and Durlauf (2000b) contain some results on computing bounds on interactions effects using Manski's methodology.

affluent communities apparently had not done so, at least by the time data were collected. Hence, serious self-selection problems exist with the data which taken as a whole make any extrapolation of the asserted findings of neighborhood effects from the program to more systematic policy interventions possibly spurious. Second, the findings of differential outcomes associated with residence in richer versus poorer neighborhoods cannot be interpreted causally without better efforts to control for individual-level effects which may vary systematically with the neighborhoods. For example, the finding that there is a reduction in asthma attacks among MTO children who move to better neighborhoods (Katz, Kling, and Liebman (1999)) is not compelling evidence of neighborhood effects without adequate controls for differences in housing quality across neighborhoods.[20] Third, the MTO data are associated with a sufficiently short time horizon that one must be worried about the distinction between transitory and permanent effects. It is far from inconceivable that there could, over time, be regression in the better outcomes associated with the movement of some families to richer communities. My own judgment is that the MTO demonstration complements the Gautreaux data, but certainly does not supplant it.

***iv*. observational data studies**

A vast number of studies using observational data have looked for interaction effects. One primary area of such research has focused on the role of neighborhood effects on children. Letting $t-1$ denote youth and t denote adulthood, in a typical specification, some outcome $\omega_{i,t}$ is related to a set of individual-specific controls and a set of neighborhood-level controls. Using our previous notation for these controls, a typical equation in this literature is

$$\omega_{i,t} = k + c'X_{i,t-1} + d'Y_{g(i),t-1} + \epsilon_{i,t}. \tag{44}$$

[20]This is more than an idle criticism. Yinger (2000), for example argues that housing discrimination exacerbates the exposure of the poor to health risks due to housing, stating that "Living in lower-quality housing implies a higher risk of exposure to lead paint or asthma inducing conditions" (pg. 27). Hence, the question of whether housing quality differs for the "treatment" and "control" groups in the MTO demonstrations is a serious one in interpreting the asserted health findings in various studies.

Well known examples of this type of study include Brooks-Gunn, Duncan, Klebanov, and Sealand (1993) and Corcoran, Gordon, Laren, G. Solon (1992). Interesting recent examples include Weinberg, Reagan and Yankow (1999) and Zax and Rees (1998). Binary choice analogs to this model have been studied by Crane (1991) for the cases of high school dropout rates and teenage fertility. Longitudinal analogs have been studied for the cases of onset of sexual activity (Brewster (1994a,b)), teenage fertility (Sucoff and Upchurch (1998)) and nonmarital fertility (South and Crowder (1999)).

These studies have been subjected to several criticisms which mitigate against the strength of their support for an interactions-based perspective. One problem is that since neighborhoods are endogenous, these types of exercises are subject to selection biases. The intuition is simple. If there are omitted individual characteristics, and these characteristics in turn help determine neighborhood memberships, then the error term $\epsilon_{i,t}$ will presumably be correlated with $Y_{g(i),t-1}$. Evidence that these sorts of self-selection problems are important has been provided by Evans, Oates, and Schwab (1992). In this study, the use of instrumental variables to control for the endogeneity of neighborhoods causes initial evidence of neighborhood effects to disappear.

Second, as emphasized by Manski (1993), empirical work of this type has generally not been precise in the modelling of interaction effects. For example, eq. (44) fails to allow for any endogenous effects. Further, this equation cannot be justified as the reduced form for a structural model which contains both endogenous and contextual effects. To see this, suppose that the structural model for individual behavior is

$$\omega_{i,t} = k + c'X_{i,t-1} + d'Y_{g(i),t-1} + Jm_{g(i),t} + \epsilon_{i,t}. \tag{45}$$

where expectations are conditional on $X_{j,t-1}$, $\forall j \in g(i)$, and $Y_{g(i),t-1}$. Taking expected values of both sides of this equation yields

$$m_{g(i),t} = k + c'X_{g(i),t-1} + d'Y_{g(i),t-1} + Jm_{g(i),t} \tag{46}$$

where, as before, $X_{g(i),t-1}$ is the sample average of $X_{j,t-1}$ for $j \in g(i)$. This expression can be rearranged to produce

$$m_{g(i),t} = \frac{k}{1-J} + \frac{c'}{1-J} X_{g(i),t-1} + \frac{d'}{1-J} Y_{g(i),t-1} \tag{47}$$

so that

$$\omega_{i,t} = \frac{k}{1-J} + c' X_{i,t-1} + \frac{c'}{1-J} X_{g(i),t-1} + \frac{d'}{1-J} Y_{g(i),t-1} + \epsilon_{i,t}. \tag{48}$$

which fails to parallel eq. (44) due to the presence of $X_{g(i),t-1}$ in (48). A useful exercise would be to reexamine those empirical studies which have neglected endogenous effects, formulate proper reduced form regressions, and see what structural parameters may be recovered. As discussed earlier, the identification of these structural parameters will depend on prior information concerning the relationship between $X_{g(i),t-1}$ and $Y_{g(i),t-1}$.

7. Policy

An interactions-based perspective on socioeconomic outcomes has important implications for the design and evaluation of public policy. How do interactions-based models affect the assessment of public policies? There are several dimensions along which one may explore this question.

From the perspective of policy assessment, interactions-based models make clear the importance of accounting for nonlinearities. To see this, suppose that a policymaker is assessing the effects of altering h, the private incentive of each individual within a population whose equilibrium choice level is described by eq. (22). The derivative of the equilibrium expected average choice m with respect to h is, for all values of βJ

$$\frac{dm}{dh} = \frac{\beta(1 - tanh^2(\beta h + \beta J m))}{(1 - \beta J(1 - tanh^2(\beta h + \beta J m)))} \tag{49}$$

if one assumes that when there are multiple equilibria, the equilibrium which is prevailing is unaffected by dh. This is obviously highly nonlinear and the effects of the change in h will be sensitive to the initial state of the population. For example, suppose that $h > 0$. If the equilibrium m is unique, one can verify that the marginal effect on a group will be higher the weaker the group's fundamentals. In the case where there are multiple equilibria, then the case is more complicated. If $h > 0$ and $m > 0$, then it is still the case that the marginal effect of a change in h is decreasing in the level of h. On the other hand, if $h > 0$ and $m < 0$, the opposite is true – higher initial levels of h imply higher marginal effects. Further, because of the relationship between fundamentals and the number of equilibria, one must additionally ask whether nonmarginal changes in h will alter the number of equilibria. This creates the possibility that it is cost-efficient to raise the private incentives for the relatively better off. For example, in Figure 1, when one moves from low to high h, the welfare inferior equilibria disappear. The bottom line from these considerations is that policy effects will not have simple multipliers of the types found in linear models.

These nonlinearities have important implications for the allocation of resources by a policymaker. Suppose that a policymaker is considering whether to implement a system of subsidies to raise incentives for I individuals, each with an identical h, out of a population of I^2. Suppose these individuals are organized into I separate groups; as before assume each group's equilibrium is described by (22). The policymaker is assumed to have two options: 1) raise incentives from h to $h + dh$ for I individuals scattered across I different groups, or 2) raise incentives from h to $h + dh$ for all I members of a given group. Let m_1 denote the expected value of the average choice for I persons sampled across the I groups and m_2 denote the expected value of the average choice for the I members of a given group. Assuming that the groups are large, so that one can ignore the effect of the behavior of one individual on the group, then the total effect on behavior under the first policy option is

$$\frac{dm_1}{dh} = \beta(1 - tanh^2(\beta h + \beta Jm)). \tag{50}$$

Eq. (49) gives the effect of the second policy. Therefore, the relative impact of the the two policies is

$$\frac{dm_1}{dh}/\frac{dm_2}{dh} = \frac{1}{(1 - \beta J(1 - tanh^2(\beta h + \beta Jm)))} \tag{51}$$

This is the social multiplier generated by the presence of social interaction effects. Intuitively, the interactions within a group magnify the effects of the private incentive changes. This means that a cost-benefit analysis would suggest concentrating expenditures in order to take advantage of the social multiplier which amplifies the effects of a higher h within a given group.

Further, interactions-based models suggest the importance of exploring alternatives to forms of redistribution which are designed to raise private incentives. One way to interpret welfare and other cash and/or in-kind aid programs is that they are forms of income redistribution. Such programs typically transfer (through taxes paid either contemporaneously or over time to retire government debt which funded the initial program) income from one group to another. An alternative form of equality-enhancing policies fall under the category of "associational redistribution" (Durlauf (1996c)). These policies treat group memberships as potential objects of redistribution.

A number of past and current public policies are interpretable as promoting associational redistribution. For example, many education policies are attempts to engage in associational redistribution. Affirmative action in college admissions is nothing more than a choice of what criteria are used to construct student bodies. School busing for racial integration, while substantially less important now than 20 years ago, had exactly the same effect. Recent efforts to promote integration through magnet schools may be interpreted the same way.

Associational redistribution is a far more controversial class of policies than standard tax/transfer policies, as the visceral public hostility to affirmative action makes clear. Further, it seems clear that the development of a rigorous ethical defense of associational redistribution is more difficult than for income redistribution; even as egalitarian a thinker as Walzer (1983) finds various forms of quotas to be ethically problematic. While the presence of interactions in determining socioeconomic outcomes cannot, of course, resolve these complexities, their presence is nevertheless important in assessing whether particular forms of associational redistribution are just.

For example, suppose one follows Roemer (1998) and concludes that society ought to indemnify individuals against adverse outcomes in life to the extent the outcomes are caused by factors outside their control. Clearly, ethnicity, residential neighborhood of youth, and the like are not variables which one chooses. Hence, the pursuit of equality of opportunity along the lines outlined by Roemer would require interventions to render these groupings irrelevant in predicting socioeconomic outcomes. One obvious way and perhaps necessary way to achieve this is to alter those group memberships which are not immutable.

8. Conclusions

This paper has described a general model of social interactions which attempts to combine the rigorous choice-based modelling of economics with the richer social structures, interdependences, and contexts, which are the hallmark of sociology. The theoretical framework embodies methodological individualism, yet illustrates how social context breaks the reduction of aggregate behavior to individual level descriptions. In terms of conceptualizing behaviors, the approach allows one to integrate private incentives and social influences in a common structure. This framework is compatible with structural econometric analysis and so can be falsified using standard statistical methods.

In terms of future research, my own view is that the most important contributions can be made in the areas of statistical methodology and empirical work. As the survey of evidence suggests, the strongest evidence in favor of social interactions lies in those contexts most removed from the substantive phenomena that this new literature tries to address. Advances in this regard will probably require much more attention to data collection. For example, virtually no attention has been paid to the question of identifying which groups influence individuals as opposed to which groups are currently measured; as Manski (1993) argues, identification of relevant groups from data is probably impossible. Census tracts may have been chosen to approximate homogeneous families, but there is no reason to believe they actually define reference groups. This suggests that detailed survey information may be needed to elicit information on what groups actually matter to individuals in a sample. In terms of policy interventions such as MTO, where one can identify the effects of changing some group characteristics, it seems clear that greater attention to experimental design and the full range of forms of self-selection is critical.

At a minimum, the framework helps to make clear, I believe, that the disciplinary barriers between sociology and economics are in many respects artificial. For phenomena such as inner city poverty, social pathologies, and the like, each field contains important and fundamental theoretical ideas that not only are compatible but are in fact complementary to one another. My own belief is that the continuing synthesis of choice-based reasoning with social interactions will prove to be one of the most promising areas of *socioeconomic* theory and empirical work.

Technical Appendix

Identification in the binary choice model with social interactions

For the binary choice model with social interactions such that

$$\mu(\underset{\sim}{\omega} \mid X_i, Y_{g(i)}, m^e_{g(i)} \ \forall\ i) \propto$$

$$\prod_i (exp(\beta k\omega_i + \beta c' X_i \omega_i + \beta d' \underset{\sim}{Y}_{g(i)} \omega_i + \beta J m^e_{g(i)} \omega_i)) \tag{A.1}$$

such that

$$m^e_{g(i)} = m_{g(i)} = \int tanh(\beta k + \beta c' \underset{\sim}{X} + \beta d' \underset{\sim}{Y}_{g(i)} + \beta J m_{g(i)}) dF_{\underset{\sim}{X} \mid \underset{\sim}{Y}_{g(i)}} \tag{A.2}$$

and β is normalized to 1, if

i. $supp(X_i, Y_{g(i)})$ is not contained in a proper linear subspace of R^{r+s}.

ii. $supp(Y_{g(i)})$ is not contained in a proper linear subspace of R^s.

iii. No element of X_i or $Y_{g(i)}$ is constant.

iv. There exists at least one group g_0 such that conditional on $\underset{\sim}{Y}_{g_0}$, $\underset{\sim}{X}_i$ is not contained in a proper linear subspace of R^r.

v. None of the regressors in $\underset{\sim}{Y}_{g(i)}$ possesses bounded support.

vi. $m_{g(i)}$ is not constant across all groups g.

Then, (k,c,d,J) is identified relative to any distinct alternative $(\overline{k}, \overline{c}, \overline{d}, \overline{J})$.

Proof: See Brock and Durlauf (2000b).

Bibliography

Abbott, A., (1988), "Transcending General Linear Reality," Sociological Theory, 6, 169-186.

Abbott, A., (1997), "Of Time and Space: The Contemporary Relevance of the Chicago School," Social Forces, 75, 1149-1182.

Akerlof, G., (1997), "Social Distance and Social Decisions," Econometrica, 65, 1005-1028.

Akerlof, G. and R. Kranton, (1999), "Identity and Economics," mimeo, University of California at Berkeley and forthcoming, Quarterly Journal of Economics.

Anderson, E., (1999), Code of the Street, New York: W. W. Norton and Company.

Anderson, P., (1972), "More is Different," Science, 177, 393-396.

Anderson, P. and D. Stein, (1984), "Broken Symmetry, Emergent Properties, Dissipative Structures, Life: Are They All Related," in Basic Notions of Condensed Matter Physics, P. Anderson, ed., Menlo Park: Addison Wesley.

Aronson, E., (1999), The Social Animal, 8th Edition, New York: W. H. Freeman and Company.

Arthur, W. B., (1987), "Urban Systems and Historical Path Dependence," in Urban Systems and Infrastructure, R. Herman and J. Ausubel, eds., Washington D.C.: National Academy of Sciences/National Academy of Engineering.

Ascher, A., (1988), The Revolution of 1905: Russia in Disarray, Stanford: Stanford University Press.

Ascher, A., (1992), The Revolution of 1905: Authority Restored, Stanford: Stanford University Press.

Bénabou, R., (1993), "Workings of a City: Location, Education, and Production," Quarterly Journal of Economics, CVIII, 619-652, 1993.

Bénabou, R., (1996a), "Equity and Efficiency in Human Capital Investment: The Local Connection," Review of Economic Studies, 62, 237-264.

Bénabou, R., (1996b), "Heterogeneity, Stratification, and Growth: Macroeconomic Effects of Community Structure," American Economic Review 86, 584-609.

Bikhchandani, S., D. Hirshleifer, and I. Welch, (1992), "A Theory of Fads, Fashion, Custom, and Cultural Exchange as Information Cascades," Journal of Political Economy, 100, 992-1026.

Blalock, H., (1984), "Contextual-Effects Models: Theoretical and Methodological Issues," Annual Review of Sociology, 10, 353-372.

Blume, L. and S. Durlauf, (1999), "The Interactions-Based Approach to Socioeconomic Behavior," mimeo, University of Wisconsin and forthcoming, Social Dynamics, H. P, Young and S. Durlauf, eds., Cambridge: MIT Press.

Blume, L. and S. Durlauf, (2000), "Equilibrium Concepts for Social Interactions Models," mimeo, Cornell University.

Borjas, G., (1992), "Ethnic Capital and Intergenerational Income Mobility," Quarterly Journal of Economics, CVII, 123-150.

Borjas, G., (1995), "Ethnicity, Neighborhoods, and Human-Capital Externalities," American Economic Review, 85, 365-390.

Brewster, K., (1994a), "Race Differences in Sexual Activity Among Adolescent Women: The Role of Neighborhood Characteristics," American Sociological Review, 59, 408-424.

Brewster, K., (1994b), "Neighborhood Context and the Transition to Sexual Activity Among Black Women," Demography, 31, 603-614.

Brock, W. and S. Durlauf (2000a), "Discrete Choice with Social Interactions" mimeo, Department of Economics, University of Wisconsin and forthcoming, Review of Economic Studies.

Brock, W. and S. Durlauf (2000b), "Interactions-Based Models," mimeo, University of Wisconsin and forthcoming, Handbook of Econometrics 5, J. Heckman and E. Leamer, eds., Amsterdam: North-Holland.

Brooks-Gunn, J., G. Duncan, P. Klebanov, and N. Sealand, (1993), "Do Neighborhoods Affect Child and Adolescent Development?," American Journal of Sociology, 99, 353-395.

Browning, C., (1992), Ordinary Men, New York: Harper Collins.

Card, D. and A. Krueger, (1992), "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," Journal of Political Economy, 1-40.

Case, A. and L. Katz (1991), "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Families," Working Paper no. 3705, National Bureau of Economics Research, Cambridge.

Chambers, J. K., (1995), Sociolinguistic Theory, Oxford: Blackwell Publishers.

Coleman, J., (1990), Foundations of Social Theory, Cambridge: Harvard University Press.

Cooper, R., (1997), Coordination Games, Cambridge: Cambridge University Press.

Corcoran, M., R. Gordon, D. Laren, and G. Solon (1992), "The Association Men's Economic Status and Their Family and Community Origins," Journal of Human Resources, 27, 575-601.

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

Currie, J. and D. Thomas, (1995), "Does Head Start Make a Difference?," American Economic Review, 85, 341-364.

Datcher, L., (1982), "Effects of Community and Family Background on Achievement," Review of Economics and Statistics, 64, 32-41.

Duneier, M., (1992), Slim's Table, Chicago: University of Chicago Press.

Duneier, M., (1999), Sidewalk, New York: Farrar, Straus, and Giroux.

Durlauf, S., (1996a), "A Theory of Persistent Income Inequality," Journal of Economic Growth, 1, 75-93.

Durlauf, S., (1996b), "Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality," in Dynamic Disequilibrium Modelling, W. Barnett, G. Gandolfo, and C. Hillinger, eds., Cambridge University Press.

Durlauf, S., (1996c), "Associational Redistribution: A Defense," Politics and Society, 24, 391-410.

Durlauf, S., (1997a), "Statistical Mechanics Approaches to Socioeconomic Behavior," in The Economy as a Complex Evolving System II, W. B. Arthur, S. Durlauf and D. Lane, eds. Redwood City: Addison-Wesley, forthcoming.

Durlauf, S., (1999), "The Memberships Theory of Inequality: Ideas and Implications," in Elites, Minorities, and Economic Growth, E. Brezis and P. Temin, eds., North Holland.

Elster, J., (1989), The Cement of Society, Cambridge: Cambridge University Press.

Evans, W., W. Oates, and R. Schwab, (1992), "Measuring Peer Group Effects: A Study of Teenage Behavior," Journal of Political Economy, 100, 966-991.

Figes, O., (1997), A People's Tragedy: The Russian Revolution 1891-1924, New York: Viking Books.

Finley, M., (1985), Democracies Ancient and Modern, second edition, New Brunswick: Rutgers University Press.

Freedman, D., (1987), "As Others See Us: A Case Study in Path Analysis," Journal of Educational Statistics, 12, 101-223.

Freedman, D., (1991), "Statistical Models and Shoe Leather," in Sociological Methodology, P. Marsden, ed., Cambridge: Basil Blackwell.

Glaeser, E., B. Sacerdote, and J. Scheinkman, (1996), "Crime and Social Interactions," Quarterly Journal of Economics, CXI, 507-548.

Glaeser, E. and J. Scheinkman, (2000a), "Measuring Social Interactions," mimeo, Princeton University and forthcoming, Social Dynamics, S. Durlauf and H. P. Young, eds., Cambridge: MIT Press.

Glaeser, E. and J. Scheinkman, (2000b), "Nonmarket Interactions," mimeo, Princeton University.

Goering, J., (1996), "Expanding Housing Choices for HUD-Assisted Families: First Biennial Report on The Moving to Opportunity for Fair Housing Demonstration," Department of Housing and Urban Development, Office of Policy Development and Research.

Granovetter, M. and R. Soong, (1988), "Threshold Models of Diversity: Chinese Restaurants, Residential Segregation, and the Spiral of Silence," in Sociological Methodology, vol. 18, C. Clogg, ed., Cambridge: Basil Blackwell.

Hanushek, E., (1986), "The Economics of Schooling: Production and Efficiency in Public Schools," Journal of Economic Literature, 24, 1141-1177.

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

Haveman, R. and B. Wolfe, (1994), Succeeding Generations, New York: Russell Sage Foundation.

Heckman, J., (1979), "Sample Selection Bias as a Specification Error," Econometrica, 47, 153-161.

Heckman, J., (1997), "Instrumental Variables," Journal of Human Resources, 32, 441-462.

Heckman, J., (2000), "Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective," Quarterly Journal of Economics, 45-97.

Heckman, J., A. Layne-Farrar, and P. Todd, (1996), "Does Measured School Quality Really Matter? An Examination of the Earnings-Quality Relationship," in Does Money Matter?, G. Burtless, ed., Washington DC: Brookings Institution Press.

Henderson, J V., P. Mieszkowski, and Y. Sauvageau, (1978), "Peer Group Effects and

Education Production Functions," Journal of Public Economics, 10, 97-106.

Horowitz, J., (1998), Semiparametric Methods in Econometrics, New York: Springer-Verlag.

Jencks, C. and S. Mayer, (1990), "The Social Consequences of Growing Up in a Poor Neighborhood," in Inner-City Poverty in the United States, L. Lynn and M. McGreary, eds., Washington D.C.: National Academy Press.

Katz, L., J. Kling and J. Liebman, (1997), Moving to Opportunity in Boston: Early Impacts of a Housing Mobility Program, mimeo, Harvard University.

Kozol, J., (1991), Savage Inequalities, New York: Crown Publishers.

Krugman, P., (1996), The Self-Organizing Economy, Oxford: Basil Blackwell.

Ladd, H. and J. Ludwig, (1998), "The Effects of MTO on Educational Opportunities in Baltimore: Early Evidence," working paper, Joint Center for Poverty Research, Northwestern University.

Loury, G., (1977), "A Dynamic Theory of Racial Income Differences," in Women, Minorities, and Employment Discrimination, P. Wallace and A. LaMond, eds., Lexington: Lexington Books.

Lucas, R., (1988), "On the Mechanics of Economic Development," Journal of Monetary Economics, 22, 3-42.

Manski, C., (1988), "Identification of Binary Response Models," Journal of the American Statistical Association, 83, 729-738.

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

Manski, C., (2000), "Economic Analysis of Social Interactions", forthcoming, Journal of Economic Perspectives.

Massey, D. and S. Kanaiaupuni, (1993), "Public Housing and the Concentration of Poverty," Social Science Quarterly, 74, 109-122.

McManus, D., (1992), "How Common is Identification in Parametric Models?," Journal of Econometrics, 53, 5-23.

Moffitt, R., (1998), "Policy Interventions, Low-Level Equilibria, and Social Interactions," mimeo, Johns Hopkins University and forthcoming, Social Dynamics, S. Durlauf and H. P. Young, eds., Cambridge: MIT Press.

Montgomery, J., (1991), "Social Networks and Labor-Market Analysis: Towards an Economic Analysis," American Economic Review, 81, 1408-1418.

Montgomery, J., (1992), "Social Networks and Persistent Inequality in Labor Markets," mimeo, Northwestern University.

Ober, J., (1989), Mass and Elite in Democratic Athens, Princeton: Princeton University Press.

Page, M. and G. Solon, (2000), "Correlations Between Sisters and Neighboring Girls in their Subsequent Income as Adults," mimeo, University of Michigan.

Popkin, S., J. Rosenbaum, and P. Meaden, (1993), "Labor Market Experiences of Low-Income Black Women in Middle-Class Suburbs: Evidence from a Survey of Gautreaux Program Participants," Journal of Policy Analysis and Management, 12, 556-574.

Portes, A., (1998), "Social Capital: Its Origins and Application in Modern Sociology," Annual Review of Sociology, 1-14.

Robertson, P., (1952), Revolutions of 1848, Princeton: Princeton University Press.

Roemer, J., (1998), Equality of Opportunity, Cambridge: Harvard University Press.

Romer, P., (1986), "Increasing Returns and Long Run Growth," Journal of Political Economy, 94, 1002-1037.

Rosenbaum, J., (1995), "Changing the Geography of Opportunity by Expanding Residential Choice: Lessons from the Gautreaux Program," Housing Policy Debate, 6, 231-269.

Rosenbaum, J. and S. Popkin, (1991), "Employment and Earnings of Low-Income Blacks Who Move to Middle Class Suburbs," in The Urban Underclass, C. Jencks and P. Peterson eds., Washington DC: Brookings Institution Press.

Rosenbaum, J., S. DeLuca, and S. Miller, (1999), "The Long-Term Effects of Residential Mobility on AFDC Receipt: Studying the Gautreaux Program with Administrative Data," mimeo, Northwestern University.

Rubinstein, A. (1998), Modeling Bounded Rationality, Cambridge: MIT Press.

Rubinowitz, L. and J. Rosenbaum, (2000), Crossing the Class and Color Lines, Chicago: University of Chicago Press.

Sampson, R., J. Morenoff, and F. Earls, (1999), "Beyond Social Capital: Collective Efficacy for Children," American Sociological Review, 64, 633-660.

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

Samuelson, L., (1997), Evolutionary Games and Equilibrium Selection, Cambridge: MIT Press.

Sampson, R., S. Raudenbush, and F. Earls, (1997), "Neighborhoods and Violent Crime: A Study of Collective Efficacy," Science, 277, 918-924.

Sobel, M., (1998), "Causal Inference in Statistical Models of the Process of Socioeconomic Achievement: A Case Study," Sociological Methods and Research, 27, 318-348.

Solow, R., (1956), "A Contribution to the Theory of Economic Growth," Quarterly Journal of Economics, LXX, 65-94.

South, S. and K. Crowder, (1999), "Neighborhood Effects on Family Formation: Concentrated Poverty and Beyond," American Sociological Review, 64, 113-132.

Sperber, J., (1994), The European Revolutions 1848-1851, Cambridge: Cambridge University Press.

Streufert, P., (1991), "The Effect of Underclass Isolation on Schooling Choice," mimeo, University of Wisconsin at Madison.

Sucoff, C. and D. Upchurch, (1998), "Neighborhood Context and the Risk of Childbearing Among Metropolitan-Area Black Adolescents," American Sociological Review, 63, 571-585.

Topa, G., (1999), "Social Interactions, Local Spillovers and Unemployment," mimeo, New York University and forthcoming, Review of Economic Studies.

Wallis, K., (1980), "Econometric Implications of the Rational Expectations Hypothesis," Econometrica, 48, 49-73.

Walzer, M. (1983), Spheres of Justice, New York: Basic Books.

Weinberg, B., P. Reagan, and J. Yankow, (1999), "Do Neighborhoods Affect Work Behavior? Evidence from the NLSY79," mimeo, Department of Economics, Ohio State University.

Wilson, W. J., (1987), The Truly Disadvantaged, Chicago: University of Chicago Press.

Yeomans, J., (1992), Statistical Mechanics of Phase Transitions, Oxford: Oxford University Press.

Yinger, J., (2000), "Housing Discrimination and Residential Segregation as Causes of Poverty," mimeo, Syracuse University.

Zax, J., and D. Rees, (1998), "Environment, Ability, Effort, and Earnings," working paper, no. 9801, Center for Research on Economic and Social Policy, University of Colorado at Denver.