Discrete Choice with Social Interactions

William A. Brock

Department of Economics, University of Wisconsin at Madison

Steven N. Durlauf

Department of Economics, University of Wisconsin at Madison

Final Revision: May 15, 2000

Forthcoming, Review of Economic Studies. The Vilas Trust, Romnes Trust, National Science Foundation, and John D. and Catherine T. MacArthur Foundation have generously provided financial support. We thank Andros Kourtellos, Artur Minkin and Eldar Nigmatullin for research assistance and seminar participants at Cambridge, Hebrew University, London School of Economics, National Bureau of Economic Research, New York University, Santa Fe Institute, UC Davis, and University College London for helpful suggestions. We are especially grateful to Orazio Attanasio, Ann Bell, Kim-Sau Chung, James Heckman, Charles Manski and two anonymous referees for providing valuable comments.

Discrete Choice with Social Interactions

Abstract

This paper provides an analysis of aggregate behavioral outcomes when individual utility exhibits social interaction effects. We study generalized logistic models of individual choice which incorporate terms reflecting the desire of individuals to conform to the behavior of others in an environment of noncooperative decisionmaking. Laws of large numbers are generated in such environments. Multiplicity of equilibria in these models, which are equivalent to the existence of multiple self-consistent means for average choice behavior, will exist when the social interactions exceed a particular threshold. Local stability of these multiple equilibria is also studied. The properties of the noncooperative economy are contrasted with the properties of an economy in which a social planner determines the set of individual choices. Finally, likelihood function based on the theoretical model is given and conditions for the econometric identifiability of the model is established.

William A. Brock
Department of Economics
1180 Observatory Drive
University of Wisconsin
Madison, WI 53706-1393
email: wbrock@ssc.wisc.edu

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

1. Introduction

A large body of recent research has begun to consider the role of social interactions in economic behavior. By social interactions, we refer to the idea that the utility or payoff an individual receives from a given action depends directly on the choices of others in that individual's reference group, as opposed to the sort of dependence which occurs through the intermediation of markets. This type of spillover is an example of a classical nonpecuniary externality (cf. Arrow and Hahn (1971), chapter 6). When these spillovers are positive in the sense that the payoff for a particular action is higher for one agent when others behave similarly, the presence of social interactions will induce a tendency for conformity in behavior across members of a reference group. Further, as described by Bernheim (1994), even when the underlying intrinsic utility from the actions differs widely across individuals due to heterogeneity of individual characteristics, the presence of this desire to conform may create either a tendency towards common behavior or towards a few polarized types of behavior within a reference group. In addition, social interactions can also represent an explanation for large cross-group variations in behavior if different groups conform to alternative types of self-reinforcing behavior. When social interactions act as strategic complementarities between agents, multiple equilibria may occur in absence of any coordination mechanisms, as described by Cooper and John (1988).

The intuition that individuals seek to conform to the behavior of reference groups has found successful application in a number of circumstances.[1] One important case is the effect of social interactions which occur within a neighborhood. Bénabou (1993) shows that when the cost of individual education investment is a decreasing function of the investment decisions of one's neighbors,

[1]A very early study of the role of social interactions in binary choice is Schelling (1973), who provides a wealth of charming examples, ranging from driving patterns to styles of athletic play.

neighborhoods can exhibit substantial discrepancies in the level of human capital formation. One interpretation of this spillover is that deviation from a neighborhood's mean education level is costly. This type of spillover can have powerful consequences for income distribution; in fact as shown by Loury (1977) in the context of racial groups and Durlauf (1996a,b) in the context of income classes such effects, when intergenerational, can lead to permanent inequality between family dynasties. Alternatively, Schelling (1971) shows how preferences over neighborhood racial composition can lead to pronounced residential segregation, even when these preferences are relatively weak.

The emphasis on social interactions as a determinant of behavior, while relatively recent in the context of economic theorizing, is of course not new from the perspective of sociology. One early analysis of the role of social interactions is found in the literature on the "culture of poverty" (see Lewis (1966) and Liebow (1967) for classic formulations and Montgomery (1994) for an interesting formalization), which argues that isolated poor groups exhibit different values towards work, childbearing and parenting from the population as a whole. More recent treatments of ghetto poverty, such as Wilson (1987), even while rejecting a strict cultural explanation for phenomena such as labor force withdrawal and out-of-wedlock births, nonetheless do emphasize the social multiplier which converts changes in private utility to changes in community-wide behavior. In fact, this interdependence between private incentives and imitative behavior will drive our theoretical framework. Our work is also closely related to Becker (1996) which stresses the role of social capital as well as individual-specific capital in explaining behavior.

Recent empirical evidence which is consistent with the presence of social interaction effects has been developed in a number of contexts. In terms of neighborhood influences, Case and Katz (1991) provide evidence that the probability of social ills in one neighborhood is increasing in the prevalence of these same ills in adjacent neighborhoods. Crane (1991) finds a relationship between both school dropout and teenage childbearing rates and the occupational composition of a community. Haveman and Wolfe (1994) present similar findings, concluding in terms of high school dropout rates, for example, that "If those who grew up in a "bad" neighborhood were to grow up in a "good" neighborhood, (the) probability of dropping out falls by 52%" (pg. 250). Similarly, Glaeser, Sacerdote, and Scheinkman (1996) argue that social interactions can explain large differences in community crime rates. Steinberg et al (1996) provide a wide range of evidence in support of the claim that peer group effects are more important that parental influences in determining high school performance. Finally, Anderson (1990) provides a fascinating portrait of the power of social interactions on individual behavior in the inner city of Philadelphia based on direct field observation.[2]

The potential role of social interactions has also been demonstrated in economic situations far removed from neighborhoods. Brock (1993) shows how these types of effects, when embedded in the expectations formation process, can help explain asset market volatility. Brock and Hommes (1997) further show how these effects, when embedded in a model of costly learning, can produce complex aggregate price dynamics. In a very different context, Dasgupta (1995), Kohler (1997), and Durlauf and Walker (1998) argue that social interactions play a major role in explaining variations in fertility rates and the adoption of different birth control technologies.

This paper is designed to provide an analytical framework – random

[2]At the same time, three caveats should be stressed in interpreting these types of studies. First, as well described in Jencks and Mayer (1990), the evidence of these effects is often not robust across studies and regression specifications. Second, the empirical literature on the Tiebout model and local public goods has long stressed the difficulty in distinguishing between evidence of group effects and the presence of correlated (within-group) individual effects, especially when these effects are themselves among the determinants of group formation. See Moffitt (2000) and Brock and Durlauf (2000) for recent discussions. Third, as shown by Manski (1993,b), there is a separate identification problem concerning whether social interactions occur due to the influence of the behavior of an individual's peers versus the characteristics of the peers. These issues will be discussed in more detail in the econometrics section below.

fields – for studying economies in which social interactions are embedded in individual decisions.[3] Random fields modelling has proven useful in studying the potential for multiple equilibria and complex cross-section dynamics in large heterogeneous economies (see Föllmer (1974) for an early contribution, Brock (1993), Durlauf (1993,1997) and Bell (1994) for recent theoretical analyses, and Topa (1999) for an econometric analysis of social interactions in unemployment, all of which use different random field structures) in which all agent interactions are local, *i.e.* when individuals have incentives to conform to the behavior of a small number of appropriately defined neighbors. Our current analysis shows how to derive complementary conclusions when the interactions are global, *i.e.* where individuals face incentives to conform their behavior to the mean of a common reference group as well as for cases in which the population size is arbitrary. Unlike previous work in this area, we are able to derive these probabilistic interdependences through an explicit analysis of interdependent utility functions for both noncooperative equilibrium and social planner environments and do so in a way which derives from individual optimization in a natural fashion.

An essential feature of our analysis is the use of parameterizations suggested by the discrete choice literature to embody social interactions. Our analysis follows Blume (1993) and Brock (1993) in exploiting the relationship between models of discrete choice with interaction effects and a particular random fields models. This strategy leads us to consider binary decision problems for individual agents. Such a framework naturally fits a wide array of social phenomena, such as teenage pregnancy, participation in the above ground economy versus becoming a criminal, location in a city or suburb, entry or withdrawal from the labor force, staying in or dropping out of school, etc. Under the discrete choice parameterization, the model proves to have a number of interesting theoretical properties. In particular, we are able to characterize how

[3]See also Arthur (1987,1989), Ioannides (1990), Weidlich (1992), and Krugman (1995) for applications of stochastic process models to social phenomena which are very much in the spirit of the current analysis.

private and social utility interact to produce aggregate behavior. Our framework does this without any sacrifice of econometric tractability and therefore should prove useful in a variety of applications. In fact, one may view our paper as an integration of the spillovers framework originated in Debreu (1952) to the discrete choice formulation of McFadden (1984) in order to provide a way of estimating social interaction effects.

Section 2 provides a baseline model of social interaction effects. Section 3 develops a probabilistic equilibrium characterization of individual choices under the assumption that these choices are made noncooperatively. Dynamic behavior, with a focus on the stability of various steady state average choice levels, is studied as well. Welfare analysis is conducted which shows how to rank the multiple steady states, when they exist. Section 4 develops a probabilistic equilibrium characterization of individual choices in the presence of a social planner. Section 5 considers the implications of alternative formulations of social utility. Section 6 develops some of the econometrics necessary for the empirical implementation of our model. Section 7 discusses some empirical implications of the model. Section 8 provides summary and conclusions.

2. Utility maximization in the presence of social interactions

i. Modelling individual choice with private social utility

We consider the problem of individual choice in the presence of social interactions. Formally, each individual in a population of I agents must choose a binary action at some common time. Each of these binary actions is coded into ω_i, a realization with support $\{-1,1\}$. The space of all possible sets of actions by the population is the I-tuple $\underset{\sim}{\omega} = (\omega_1, ..., \omega_I)$. Finally, $\underset{\sim}{\omega}_{-i}$ denotes $(\omega_1, ..., \omega_{i-1}, \omega_{i+1}, ..., \omega_I)$, the choices of all agents other than i.

Individual utility, $V(\omega_i)$, is assumed to consist of three components.

$$V(\omega_i) = u(\omega_i) + S(\omega_i, \mu_i^e(\underset{\sim}{\omega}_{-i})) + \epsilon(\omega_i) \tag{1}$$

The term $\mu_i^e(\underset{\sim}{\omega}_{-i})$ denotes the conditional probability measure agent i places on the choices of others at the time of making his own decision. The components of total utility are threefold: $u(\omega_i)$ is the private utility associated with a choice, $S(\omega_i, \mu_i^e(\underset{\sim}{\omega}_{-i}))$ is the social utility associated with the choice, and $\epsilon(\omega_i)$ is a random utility term, independently and identically distributed across agents. Agent i knows $\epsilon(\omega_i)$ at the time of his decision.

This formulation is closely related to a number of types of social interactions which have appeared in the literature. When $S(1, \mu_i^e(\underset{\sim}{\omega}_{-i})) - S(-1, \mu_i^e(\underset{\sim}{\omega}_{-i}))$ is increasing in a rightward shift of $\mu_i^e(\underset{\sim}{\omega}_{-i})$ (where a rightward shift in a probability measure is interpreted in a stochastic dominance sense), our social utility component will exhibit the expectational analogue to increasing differences in the sense of Milgrom and Roberts (1990), pg. 1261, and will represent a version of the binary choice models with externalities studied in Schelling (1973). In the case where $S(\omega_i, \mu_i^e(\underset{\sim}{\omega}_{-i}))$ both exhibits increasing differences and depends only on

$$\overline{m}_i^e = (I-1)^{-1} \sum_{j \neq i} m_{i,j}^e, \tag{2}$$

($m_{i,j}^e$ denotes the subjective expected value from the perspective of individual i of individual j choice), social utility will exhibit the "totalistic" form of strategic complementarities studied by Cooper and John (1988).

ii. Parametric assumptions

We restrict our analysis to parametric representations of both the social utility term and the probability density of the random utility term. These assumptions will render our model both theoretically and econometrically tractable.

First, we consider forms of social utility which exhibit a strategic complementarity that is both totalistic and constant. This means that we are interested in forms of $S(\omega_i, \overline{m}_i^e)$ which have the property that

$$\frac{\partial^2 S(\omega_i, \overline{m}_i^e)}{\partial \omega_i \partial \overline{m}_i^e} = J > 0 \tag{3}$$

A constant cross-partial allows one to measure the degree of dependence across agents with a single parameter. This assumption leads us to two functional forms for social utility. The first embodies a multiplicative interaction between individual and expected average choices,

$$S(\omega_i, \overline{m}_i^e) = J\omega_i \overline{m}_i^e. \tag{4}$$

We will designate this as the proportional spillovers case, since the percentage change in individual utility from a change in the mean decision level is constant, given the individual's choice.

The second parameterization captures a pure conformity effect of the type studied by Bernheim (1994),

$$S(\omega_i, \overline{m}_i^e) = -\frac{J}{2}(\omega_i - \overline{m}_i^e)^2. \tag{5}$$

This specification penalizes deviations far from the mean more strongly than the proportional spillovers case.

To see the relationship between the two forms of social utility, we rewrite (5) as

$$-\frac{J}{2}(\omega_i - \overline{m}_i^e)^2 = J\omega_i \overline{m}_i^e - \frac{J}{2}(1 + (\overline{m}_i^e)^2), \tag{6}$$

making use of the fact that $\omega_i^2 = 1$. This form of (5) when contrasted with (4) shows that while the two social utility specifications differ in levels, they coincide on those terms which contain the individual choice variable.

Finally, we assume that the errors $\epsilon(-1)$ and $\epsilon(1)$ are independent and extreme-value distributed, so that the differences in the errors are logistically distributed,[4]

$$Prob(\epsilon(-1)-\epsilon(1)\leq x) = \frac{1}{1+exp(-\beta x)}. \tag{7}$$

This assumption produces a direct link between the theoretical model and its econometric implementation.

As the probability that ω_i takes on the value -1 rather than 1 will equal the probability that $V(-1)>V(1)$, parameterization of the probability density of $\epsilon(-1)-\epsilon(1)$ will allow explicit calculation of $Prob(\underset{\sim}{\omega})$.

3. Equilibrium properties under noncooperative decisionmaking

i. Equilibrium under proportional spillovers

We first study the behavior of the model with the proportional spillovers specification (4) under the assumption that agents act noncooperatively. Operationally, this means that each agent makes his choice given an expectation of the mean choice level which is independent of the realizations of $\epsilon(\omega_i)\ \forall\ i$. In other words, agents do not communicate or coordinate their decisions. It is standard under the extreme-value assumption for $\epsilon(\omega_i)$ that each individual choice will obey the probability

[4]See McFadden (1984) and Anderson, de Palma, and Thisse (1992) for valuable discussions of the various motivations for the logistic model as well as derivations of many standard results.

$$Prob(\omega_i) = \frac{exp(\beta(u(\omega_i) + J\omega_i \bar{m}_i^e))}{\sum_{\nu_i \in \{-1,1\}} exp(\beta(u(\nu_i) + J\nu_i \bar{m}_i^e))}. \tag{8}$$

In this probability measure, β parametrizes the extent to which the deterministic component of utility determines actual choice. As $\beta \Rightarrow \infty$, the effect of $\epsilon(\omega_i)$ on the realized choice will vanish whereas as $\beta \Rightarrow 0$, the probability that $\omega_i = -1$ (or 1) will converge to $\frac{1}{2}$, regardless of the values of the private and social utility terms under each choice.

Since the $\epsilon(\omega_i)$ terms are independent across agents, the joint probability measure over all choices equals

$$Prob(\underset{\sim}{\omega}) = \frac{exp(\beta(\sum_{i=1}^{I}(u(\omega_i) + J\omega_i \bar{m}_i^e)))}{\sum_{\nu_1 \in \{-1,1\}} \cdots \sum_{\nu_I \in \{-1,1\}} exp(\beta(\sum_{i=1}^{I}(u(\nu_i) + J\nu_i \bar{m}_i^e)))}. \tag{9}$$

When $J = 0$, this expression is proportional to the standard logistic density; the standard logistic form follows directly when one performs the change of variables $\kappa_i = \frac{\omega_i + 1}{2}$ in order to shift the support of the individual decisions from $\{-1,1\}$ to $\{0,1\}$. Hence, when $J \neq 0$, this standard form is augmented by social interactions and this parameter may be used to fully characterize the effects of interactions on community behavior. This model possesses a probability structure equivalent to the so-called mean field version of the Curie-Weiss model of statistical mechanics, (comprehensively developed in Ellis (1985)). The properties of this joint probability measure may be developed as follows.

First, we convert eq. (9) so that the exponent in the expression only depends on ω_i linearly. This may be done as follows. Since we are dealing with binary choices, private utility $u(\omega_i)$ can be replaced with $h\omega_i + k$ where h and k are chosen so that $h + k = u(1)$ and $-h + k = u(-1)$. Notice that this implies that $h = \frac{1}{2}(u(1) - u(-1))$ and so this parameter is proportional to the deterministic private utility difference between the two choices.

Using this linearization, observe that for each individual, the expected value of his choice, conditional on his beliefs concerning the behavior of others, may be written reintroducing the expectations of individual choices as described in eq. (2)

$$\mathrm{E}(\omega_i) =$$

$$1 \cdot \frac{exp(\beta h + \beta J(I-1)^{-1}\sum_{j \neq i} m^e_{i,j})}{exp(\beta h + \beta J(I-1)^{-1}\sum_{j \neq i} m^e_{i,j}) + exp(-\beta h - \beta J(I-1)^{-1}\sum_{j \neq i} m^e_{i,j})}$$

$$-1 \cdot \frac{exp(-\beta h - \beta J(I-1)^{-1}\sum_{j \neq i} m^e_{i,j})}{exp(\beta h + \beta J(I-1)^{-1}\sum_{j \neq i} m^e_{i,j}) + exp(-\beta h - \beta J(I-1)^{-1}\sum_{j \neq i} m^e_{i,j})}$$

$$= tanh(\beta h + \beta J(I-1)^{-1}\sum_{j \neq i} m^e_{i,j}). \quad (10)$$

Finally we impose rational expectations, i.e. we require that for all i and j, $m^e_{i,j} = \mathrm{E}(\omega_j)$. Since the $tanh$ function is a continuous mapping, and since the support of the choices is $\{-1,1\}^I$, it is immediate from Brouwer's fixed point theorem that there must exist a fixed point with respect to the $\mathrm{E}(\omega_i)$'s such that

$$\mathrm{E}(\omega_i) = tanh(\beta h + \beta J(I-1)^{-1}\sum_{j \neq i} \mathrm{E}(\omega_j)). \quad (11)$$

Finally, by symmetry of these expectations equations, one may conclude that at a self-consistent equilibrium $\mathrm{E}(\omega_i) = \mathrm{E}(\omega_j)\ \forall i, j$ and that this common individual expected value must also equal the expected value of the average choice for any subset of the population. Hence a self-consistent, or rational expectations equilibrium will always exist and we conclude Proposition 1.

Proposition 1. Existence of self-consistent equilibrium for discrete choices with noncooperative decisionmaking

When agents choose actions noncooperatively given social utility specification eq. (4) and given self-consistent expectations, there exists at least one expected average choice level m^* such that

$$m^* = tanh(\beta h + \beta J m^*). \tag{12}$$

This model exhibits an analogue to multiple equilibria in deterministic models of strategic complementarities, such as those described in Cooper and John (1988). Hence, one would expect that multiple equilibria are possible; this will in fact occur when there exist multiple solutions to eq. (12). These multiple solutions imply the existence of distinct expected average choice levels which are each compatible with individually optimal decisions. Conditions for the existence of multiple solutions may be immediately obtained from the properties of the $tanh(\cdot)$ function and are summarized in Proposition 2.

Proposition 2. Existence of multiple average choice levels in equilibrium

i. If $\beta J > 1$ and $h = 0$, there exist three roots to eq. (12). One of these roots is positive, one root is zero, and one root is negative.

ii. 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 roots to eq. (12), one of which has the same sign as h, and the others possessing opposite sign.

b. for $|\beta h| > H$, there exists a unique root to eq. (12) with the same sign as h.

Proposition 2 allows us to designate without ambiguity m^*_- as the mean choice level in which the largest percentage of agents choose -1, m^*_+ as the mean choice level in which the largest percentage of agents choose 1, and m^*_m as the root associated with a mean between these two values, when there are multiple roots.

One interesting feature of the proposition is that the potential for multiple average equilibrium choice levels depends both on the strength of the social utility as well as the magnitude of the bias towards one choice induced by private utility. In other words, for each β and J, there will exist a level for h which ensures that the equilibrium is unique. This implies one is most likely to observe multiplicity in those social environments in which private utility renders individuals relatively close to indifferent between choices.

One can establish the expected percentages of positive choices in the population, $k^* = \frac{m^*+1}{2}$, when agents possess self-consistent expectations in the sense of eq. (12). The following properties for k^* are straightforward to verify.

Proposition 3. Relationship between limiting percentage of positive choices and model parameters

i. If $h = 0$ and $\beta J < 1$, $k^* = \frac{1}{2}$.

ii. $lim_{h \Rightarrow \infty}\ k^* = 1$.

iii. $lim_{h \Rightarrow -\infty}\ k^* = 0$.

iv. If $h = 0$, $lim_{J \Rightarrow \infty}\ k^* = 1$, $\frac{1}{2}$, or 0 depending on whether m^*_+, m^*_m, or m^*_- is the root of eq. (12) which characterizes the equilibrium of the economy.

v. If $h \neq 0$, then $lim_{J \Rightarrow \infty}\ k^* = 1$ or 0, depending on which root of eq. (12) characterizes the equilibrium of the economy.

***ii.* Dynamic stability**

We consider the dynamic stability of the steady equilibrium expected choice levels m^*_-, m^*_+, and m^*_m. We do this by considering the dynamics of the mean choice levels under the assumption that all variables in the original model are now subscripted by time and that the expectations term $\overline{m}^e_{i,t}$ obeys the relationship

$$\overline{m}^e_{i,t} = m^*_{t-1}\ \forall\ i \tag{13}$$

where m^*_{t-1} is the mathematical expectation of the average choice at $t-1$. In other words, we consider the dynamics of a sequence of economies in which expectations are myopic in that each agent uses the optimal forecast of last period's average choice as the basis for expectations formation. While this analysis certainly does not exhaust the analysis of learning mechanisms in the model, it does illustrate how dynamic analogs of the model will evolve.[5] Notice as well that if the spillover effects are intertemporal, as in Durlauf (1993), so that aggregate behavior last period affects current individual payoffs, our analysis will also apply.[6]

The analysis leading to eq. (12) immediately implies the existence of a unique m_t conditional on any m_{t-1}. Therefore, local stability of a particular

[5]As we shall see below, as $I \Rightarrow \infty$, the expectation and realization will coincide, so that the expectations are myopic in the standard sense that an expectation at t equal the realization of the same random variable at $t-1$.

[6]For example, the payoff to labor force participation of generation t might depend on the labor force participation decisions of generation $t-1$ due to role model or labor market connection effects.

steady state identified in Proposition 1 will require that it represents a limiting solution to

$$m_t = tanh(\beta h + \beta J m_{t-1}) \tag{14}$$

where m_0 is taken anywhere in some sufficiently small neighborhood of that steady state.

We sketch an argument on stability as follows, assuming $\beta J > 1$ and $h = 0$. In this case, the derivative of $m_t - m_{t-1}$ with respect to m_{t-1} will equal

$$\frac{\partial(m_t - m_{t-1})}{\partial m_{t-1}} = \beta J(1 - tanh^2(\beta J m_{t-1})) - 1. \tag{15}$$

Suppose we start with $m_0 = 0$. This would imply that $\frac{\partial(m_1(m_0) - m_0)}{\partial m_0} > 0$ since $tanh(0) = 0$. Hence $m^*_m = 0$ is not locally stable, since by continuity and symmetry of $tanh(\cdot)$, one could find a neighborhood around 0 such that $m_1 > m_0$ if $m_0 > 0$ and $m_1 < m_0$ if $m_0 < 0$. Now suppose that we start with $m_0 = 0^+$. By eq. (15), $m_t > m_{t-1}$ $\forall$ $t > 1$, so the sequence is monotonically increasing. Since m_t is bounded, the sequence must converge to some limit which, by Proposition 2.*i*, must be m^*_+ since there are no other steady state solutions with positive value. Therefore m^*_+ is locally stable from below. Now suppose that $m_0 = 1$. In this case, eq. (15) implies $m_1 < m_0$, eq. (14) implies the sequence is monotonically decreasing, which will again require that its limit is m^*_+ since there are no other steady states with positive value; hence m^*_+ is stable from above. This verifies local stability of m^*_+. By symmetry, m^*_- must be locally stable as well.

Analysis of the case with three roots and $|h| \neq 0$ (but small) is parallel to the $h = 0$ case. Further, analogous reasoning can be used to show the stability of the unique steady states which occur when either $\beta J < 1$ or $|h|$ is large enough. Together, this leads to Proposition 4.

Proposition 4. Stability of steady state mean choice levels

Under the assumption of noncooperative decisionmaking and the expectations formation process eq. (13),

i. If eq. (12) exhibits a unique root, that root is locally stable.

ii. If eq. (12) exhibits three roots, then the steady state mean choice levels m^*_- and m^*_+ are locally stable whereas the steady state mean choice level m^*_m is locally unstable.

This result provides an interesting complement to analyses in Miyao (1978) and Bénabou (1993) on the instability of integrated neighborhood equilibria in which there are intra-group externalities. While those analyses show how agents will segregate themselves by type in the presence of externalities, thereby inducing within-group homogeneity and cross-group heterogeneity, our analysis illustrates how agents will choose to act relatively homogeneously when their types (defined in terms of realizations of $\epsilon(\omega_i)$ across i) are heterogeneous and they are required to form a common group.

In subsequent analysis, we shall focus only on the two stable equilibria.

iii. Welfare analysis

Unlike the frameworks studied in Cooper and John (1988) and Milgrom and Roberts (1990), there will not exist a Pareto ranking across the two equilibrium mean choice levels given a realization of the individual utility errors $\epsilon_i(\omega_i)$. The reason for this is simple. Extreme realizations of these errors will cause some agents to choose -1 and others to choose 1 regardless of the social utility induced by the choices of others. Hence these agents will disagree on the relative desirability of the m^*_- and m^*_+ equilibrium means, and therefore no Pareto ranking will exist. However, one can exploit the preference symmetry across agents to calculate the expected utility of a typical agent (*i.e.* expected prior to the realization of his random utility terms) under the two equilibria, and use this to evaluate social welfare under the two mean choice levels. This calculation compares

$$\mathrm{E}(max_{\omega_i} V(\omega_i) \mid m^*_+) = \mathrm{E}max_{\omega_i}(h\omega_i + k + J\omega_i m^*_+ + \epsilon(\omega_i)) \tag{16}$$

to

$$\mathrm{E}(max_{\omega_i} V(\omega_i) \mid m^*_-) = \mathrm{E}max_{\omega_i}(h\omega_i + k + J\omega_i m^*_- + \epsilon(\omega_i)). \tag{17}$$

Anderson, de Palma and Thisse (1992) (see pg. 60-61 for a proof) show that for any root m^*, the expected utility can be written as

$$\begin{gathered}\mathrm{E}(max_{\omega_i} V(\omega_i) \mid m^*) = \\ \beta^{-1}(ln(exp(\beta h + \beta k + \beta J m^*) + exp(-\beta h + \beta k - \beta J m^*))).\end{gathered} \tag{18}$$

When $h = 0$, it is easy to show that $|m^*_+| = |m^*_-|$, so eqs. (16) and (17) must be equal. Thus in the absence of any private utility, each root provides equal expected utility. On the other hand, when $h > 0$ (< 0), $|m^*_+| > |m^*_-|$ $(|m^*_-| > |m^*_+|)$, so that the expected utility under m^*_+ (m^*_-) must exceed the expected utility under m^*_- (m^*_+). Intuitively, the root whose sign is the same as the mean that would exist in absence of any social interactions is that which maximizes expected utility since at a mean with that sign, the private and social utility effects work in the same direction. This verifies Proposition 5.

Proposition 5. Welfare rankings

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.

iv. Large economy behavior

When expectations are characterized by a self-consistent solution to (12), this implies that the probability density thus factors into the product of I independent and identically distributed Bernoulli random variables. Hence a law of large numbers will apply to the sample mean of choices within a population and we may conclude Proposition 6.

Proposition 6. Law of large numbers for discrete choices with noncooperative decisionmaking

If agents possess common self-consistent beliefs m^*, then the sample average population choice, $\bar{\omega}_I$, converges weakly to this expectation,

$$\bar{\omega}_I \Rightarrow_w m^* \tag{19}$$

where m^* is a solution of $m^* = tanh(\beta h + \beta J m^*)$.

v. Equilibrium with conformity effects

Finally, we consider the properties of a noncooperative equilibrium when social utility embodies conformity effects of the form eq. (5). In this case, via eq. (6) and given the symmetry of the individual decision problems, the joint probability for choices will obey,

$$Prob(\underset{\sim}{\omega}) =$$

$$\frac{exp(\beta(\sum_{i=1}^{I}(h\omega_i + J\omega_i \bar{m}_i^e - \frac{J}{2}(1+(\bar{m}_i^e)^2))))}{\sum_{\nu_1 \in \{-1,1\}} \cdots \sum_{\nu_I \in \{-1,1\}} exp(\beta(\sum_{i=1}^{I}(h\nu_i + J\nu_i \bar{m}_i^e - \frac{J}{2}(1+(\bar{m}_i^e)^2))))}. \tag{20}$$

Notice, though, that $exp(-\frac{\beta J}{2}(1+(\bar{m}_i^e)^2))$ cancels out of the numerator and denominator of this expression. Hence, eqs. (20) and (9) are equivalent, which means that all features we have developed for the proportional spillovers specification apply to the conformity specification as well. Further, it is straightforward to replicate the analysis and conclusions of Section 3.i under conformity effects, which leads to Proposition 7.

Proposition 7. Equivalence of noncooperative equilibrium under positive spillovers and conformity effects

Propositions 1-6 will still hold if social utility takes the form eq. (5) rather than eq. (4).

4. Choice under a social planner

In this section, we consider how a social planner would set choices when the planner's preferences are consistent with the individual utilities expressed in eq. (1) yet which preserve the analytical and econometric tractability of the discrete choice framework. We assume that a social planner possesses a utility function over the set of choices, $P(\underset{\sim}{\omega})$, which consists of deterministic and random components,

$$P(\underset{\sim}{\omega}) = U(\underset{\sim}{\omega}) + \epsilon(\underset{\sim}{\omega}). \tag{21}$$

We constrain the deterministic component of the social planner's utility to equal the sum of the deterministic components of the individual utilities in the population,

$$U(\underset{\sim}{\omega}) = \sum_{i=1}^{I} (u(\omega_i) + S(\omega_i, \underset{\sim}{\omega}_{-i})). \tag{22}$$

Notice that by placing $\underset{\sim}{\omega}_{-i}$ in the individual social utility functions, the planner internalizes the individual-level spillover effects induced by the mean choice level.

It is tempting to assume that $\epsilon(\underset{\sim}{\omega}) = \sum_i \epsilon(\omega_i)$, so that the social planner's utility is nothing more than the sum of the individual utilities. Unfortunately, this assumption would render the model analytically intractable, since the sum of a set of extreme-value distributed random variables is not extreme-value distributed. Hence, we assume that the error $\epsilon(\underset{\sim}{\omega})$ is itself independent and extreme-value distributed across all 2^I possible configurations of $\underset{\sim}{\omega}$. This assumption will ensure that the joint probability measure characterizing individual choices under a social planner has the same logistic form as the noncooperative case as discussed in Anderson, de Palma and Thisse (1992 chapter 2). By calibrating the parameters of the errors in this social planner's problem to the errors in the noncooperative problem, one can impose that the errors have the same variance. Notice that under this interpretation, the random utility of the social planner, rather than that for individuals, is germane to the determination of $\underset{\sim}{\omega}$. One interpretation of the social planner's random utility term is that it represents noise in the planner's ability to calculate tradeoffs between individual utilities. In the special case where the variance of random utility terms is zero for both individuals and the planner, the planner's preferences will correspond to the sum of the individual utilities.

This specification of a social planner determining the vector of individual choices is of interest both in terms of its contrast with the noncooperative equilibrium as well as in terms of its possible empirical relevance. As described by Coleman (1988, 1990 chapter 12), the evolution of social capital, defined to include aspects of social structure which facilitate coordination across individuals and which may be embedded either in personal mores or organizations such as churches or schools, implies that in many types of social situations, coordinated behavior can emerge. While any link between social capital and the particular planning problem we consider is heuristic, we do feel that social capital-type arguments make it important to consider population behaviors other than those which occur under noncooperation.

The analysis of the social planner's problem is complicated, and unlike the noncooperative case, analytical results are only available for large economy limits. The technical appendix, which is largely self-contained, adapts arguments found in Brock (1993) to our model.

i. Social planner problem with proportional spillovers

As before, we examine the model with proportional spillovers first. In this case, the deterministic part of the planner's utility can be written as

$$U(\underset{\sim}{\omega}) = \sum_{i=1}^{I} (u(\omega_i) + \frac{J}{I-1}(\sum_{j \neq i} \omega_j)\omega_i). \tag{23}$$

Since $\omega_i^2 = 1$, we can add and subtract $\frac{J}{I-1}$ to the right hand side and rewrite this as

$$U(\underset{\sim}{\omega}) = \sum_{i=1}^{I} (u(\omega_i) + \frac{JI}{I-1}\bar{\omega}_I \omega_i) - \frac{JI}{I-1}. \tag{24}$$

For large I, $\frac{JI}{I-1}$ becomes arbitrarily close to J and we will impose this as an approximation. The social planner's problem for our model may then be derived by replacing $\bar{m}$ with $I^{-1}\sum_{i=1}^{I} \omega_i$ in eq. (9). The probability measure characterizing the joint choice of $\underset{\sim}{\omega}$ follows the same logistic form as the noncooperative case in the sense that

$$Prob(\underset{\sim}{\omega}) = \frac{exp(\beta(\sum_{i=1}^{I} h\omega_i + \frac{J}{I}(\sum_{i=1}^{I}\omega_i)^2))}{\sum_{\nu_1 \in \{-1,1\}} \cdots \sum_{\nu_I \in \{-1,1\}} exp(\beta(\sum_{i=1}^{I} h\nu_i + \frac{J}{I}(\sum_{i=1}^{I}\nu_i)^2))}. \tag{25}$$

Unlike the noncooperative case, the likelihood of each $\underset{\sim}{\omega}$ will account for the spillover effects induced through the impact of individual choices on mean behavior, as one would expect from the eqs. (23) and (24).

In order to analyze this probability measure, which is known in the statistical mechanics literature as the Curie-Weiss model, it is necessary to eliminate the $(\sum_{i=1}^{I}\omega_i)^2$ terms in (25). This calculation is given in the technical appendix and leads to Proposition 8.

Proposition 8. Law of large numbers for individual choices in solution of social planner's problem in the presence of proportional spillovers

Let m^* denote the root of $m^* = tanh(\beta h + 2\beta J m^*)$ with the same sign as h. If eq. (25) characterizes the joint distribution of discrete choices, then

$$\overline{\omega}_I \Rightarrow_w m^*. \tag{26}$$

Recall that from Proposition 6, the equilibrium expected average choice level for the noncooperative version of this model is any solution m to $m = tanh(\beta h + \beta J m)$. This means, in the large economy limit, that there are two distinct differences between the average choice levels which occur in a noncooperative equilibrium as opposed to the case where choices are set by a social planner.

First, multiple equilibrium levels of average choice can exist without cooperation when $\beta J > 1$, whereas the average choice level is unique under the social planner. This is unsurprising given the Pareto rankability of the multiple steady states in terms of individual expected utility in the noncooperative case, when contrasted with the formulation of the social planner's utility.

Second, the average choice level chosen under the social planner's solution will be the same as would be chosen under the noncooperative solution if the signs of the means are preserved and the value of J is doubled, assuming one eliminates the multiplicity by always choosing the root of the relevant *tanh* equation whose sign is the same as h. This means that even if the mean choice level in the noncooperative equilibrium has the same sign as the social planner equilibrium, the average choice in the noncooperative case will still be socially inefficient.

Intuitively, while agents in the noncooperative equilibrium account for the effects of others on themselves, they do not account for their effect on others. One can see this by contrasting eq. (9), in which the expected value of the average choice appears in the probability measure which describes individual agents and eq. (25), in which the sample average of individual choice appears in in the probability measure which describes the social planner's collective choices. By symmetry of the spillovers across agents, as given by eq. (4), the equilibrium probability measure under noncooperative decisions ignores half of the total spillovers induced by individual decisions in the sense that while the spillovers onto individual i affect his behavior, he does not take account of the spillovers induced by his behavior. In contrast, all the spillovers are accounted for by our social planner. This failure to internalize spillover effects can, however, be offset by doubling the social interaction parameter, leading to Proposition 9.

Proposition 9. Sustainability of social planner's solution under decentralized decisionmaking with proportional spillovers

The social planner's choice of $\underset{\sim}{\omega}_I$ in the large economy limit can be supported under decentralized decisionmaking by doubling the social utility payoff to each individual, in the sense of doubling J in the noncooperative problem.

Of course, it is also possible in this case that a doubling of the social utility payoff could be counterproductive, if the equilibrium chosen noncooperatively were that which had an opposite sign to h. This is easily controlled for, in the context of a government subsidy, by ensuring that the doubling of J occurs only for choices whose sign is the same as h.

***ii.* Social planner's problem with conformity effects**

Following the same approximation as used for the proportional spillovers case, the social planner's problem with conformity effects will be characterized by the joint probability measure

$$Prob(\underset{\sim}{\omega}) = \frac{exp(\beta(\sum_{i=1}^{I} h\omega_i - \frac{J}{2}(\sum_{i=1}^{I}(\omega_i - \bar{\omega}_I)^2)))}{\sum_{\nu_1 \in \{-1,1\}} \cdots \sum_{\nu_I \in \{-1,1\}} exp(\beta(\sum_{i=1}^{I} h\nu_i - \frac{J}{2}(\sum_{i=1}^{I}(\nu_i - \bar{\nu}_I)^2)))}. \quad (27)$$

Since $-\frac{J}{2}\sum_{i=1}^{I}(\omega_i - \bar{\omega}_I)^2 = \frac{J}{2I}(\sum_{i=1}^{I}\omega_i)^2 - \frac{J}{2}I$, we can reexpress this probability as

$$Prob(\underset{\sim}{\omega}) = \frac{exp(\beta(\sum_{i=1}^{I} h\omega_i + \frac{J}{2I}(\sum_{i=1}^{I}\omega_i)^2))}{\sum_{\nu_1 \in \{-1,1\}} \cdots \sum_{\nu_I \in \{-1,1\}} exp(\beta(\sum_{i=1}^{I} h\nu_i + \frac{J}{2I}(\sum_{i=1}^{I}\nu_i)^2))} \quad (28)$$

which has the same form as (25) when J in that equation is replaced with $\frac{J}{2}$. Therefore, we can use the limiting behavior of (25) to conclude Proposition 10.

Proposition 10. Law of large numbers for discrete choices with social interactions in social planner's solution in the presence of conformity effects

Let m^* denote the root of $m^* = tanh(\beta h + \beta J m^*)$ with the same sign as h. If eq. (28) characterizes the joint distribution of discrete choices, then

$$\overline{\omega}_I \Rightarrow_w m^*. \tag{29}$$

A comparison of Proposition 10 with Proposition 1 reveals an important difference between the proportional spillovers and conformity effects specifications. In the presence of conformity effects, the mean choice level under the social planner solution is one of the steady state solutions under decentralized decisionmaking. This immediately implies Proposition 11.

Proposition 11. Sustainability of social planner's solution under decentralized decisionmaking with conformity effects

The social planner's choice of $\underset{\sim}{\omega}_I$ in the large economy limit can be supported under decentralized decisionmaking when social utility exhibits conformity effects of the form (5).

The intuition for the discrepancy between the mean choice levels in the noncooperative and social planner cases for the two social utility parameterizations may be seen by computing the expected utility of a representative agent under noncooperation, as a function of the equilibrium mean choice level. Replacing $Jm^*\omega_i$ with an arbitrary $S(\omega_i, m^*)$ in eq. (18) and differentiating with respect to m^* reveals that

$$\frac{\partial \mathrm{E}(max_{\omega_i} V(\omega_i) \mid m^*)}{\partial m^*} = \frac{\partial \mathrm{E}(S(\omega_i, m^*) \mid m^*)}{\partial m^*}. \tag{30}$$

For the proportional spillovers model, the expected utility of a representative agent with respect to his own utility innovations will, in the large economy limit of a noncooperative equilibrium, have the feature that

$$\frac{\partial \mathrm{E}(h\omega_i + k + J\omega_i m^* \mid m^*)}{\partial m^*} = Jm^* \tag{31}$$

since by self-consistency, $\mathrm{E}\omega_i = m^*$. This means that when $m^* > 0$, a marginal increase in the average choice level raises the expected utility of the typical agent whereas when $m^* < 0$, a decrease in the average choice level raises expected utility. This means, from the perspective of the social interaction component on individual utility, that there is an externality in the mean choice level which is not accounted for by individuals, as expected utility could be increased by a coordinated change in the mean.

Under conformity effects, on the other hand, the associated derivative will follow

$$\frac{\partial \mathrm{E}(h\omega_i - \frac{J}{2}(\omega_i - m^*)^2 \mid m^*)}{\partial m^*} = \mathrm{E}(J(\omega_i - m^*) \mid m^*) = 0. \tag{32}$$

Hence there is no external effect which fails to be internalized by individuals, at least locally.[7] Since (32) holds for any self-consistent m^*, it must hold at the mean choice level as determined by the social planner, whose deterministic utility component is the sum of the private deterministic utility components, so that the social planner equilibrium is sustainable under decentralized decisionmaking. Intuitively, since the conformity specification more strongly punishes large deviations from the mean than the proportional spillovers specification, the total average utility benefit from a marginal change in the mean in the direction of the

[7] In fact, it is easy to see that the proportional spillovers specification is consistent with the condition for inefficiency of a noncooperative equilibrium in Cooper-John (1988), Proposition 2, whereas the conformity specification is not. See Bryant (1983) for a similar case where efficiency can be sustained in a noncooperative environment.

majority for those in the majority is exactly offset by the utility loss to those who choose differently from the majority. Such an exact offset does not hold under proportional spillovers.

iii. Social planner's problem in absence of deterministic private utility

Finally, we consider the case $h=0$ for proportional spillovers. (The reasoning for the conformity specification is identical.) In this case, $m^*=tanh(2\beta Jm^*)$ will have three solutions if $\beta J>1$; further, no expected average welfare difference will exist between the equilibria and hence the planner will be indifferent between the two means, unlike the case of $h\neq 0$ studied in Proposition 8. Designate the two nonzero roots as y^*_+ and y^*_- and define $m^*_+=\frac{y^*_+}{2\beta J}$ and $m^*_-=\frac{y^*_-}{2\beta J}$. Ellis (1985), pg. 100, shows that the limiting probability measure over choices has the property that

$$\overline{\omega}_I \Rightarrow_w m^*_+ \text{ with probability } \tfrac{1}{2}$$

$$\overline{\omega}_I \Rightarrow_w m^*_- \text{ with probability } \tfrac{1}{2}. \tag{33}$$

In other words, the limiting measure for $\underset{\sim}{\omega}$ will be a mixture whose limiting behavior may differ across sample path realizations. This mixture has two interesting features. First, the root corresponding to $y^*=0$ does not appear in the limiting expression. This parallels the instability of this root under noncooperative decisionmaking. Intuitively, the utility from bunching means that even when spillover effects are internalized in the sense of eq. (21), the system cannot rest at $m^*=0$. Second, the probability weights on the two conditional (given y^* values) limiting means are equal. What this means is that under each sample path realization of the economy, there is an equal probability of producing the m^*_+ and m^*_- mean choice levels. Intuitively, social utility is embedded in eq. (22) in such a way that all spillovers from each individual choice are accounted for. When $h = 0$, the tendency of the mean choice level is irrelevant; what matters is that agents achieve high utility by tending to act similarly.

We conclude this section with the observation that no comparable results on the relationship between the noncooperative and social planner solutions for our model are currently available in the case where individual agents possess heterogeneous h_i's. The reason for this is that there does not currently exist any results on laws of large numbers for the Curie-Weiss model where the individual random variables are not identically distributed. Results in Amaro de Matos and Perez (1991) suggest such a generalization should be possible.

5. Extensions

We illustrate three extensions of the basic modelling framework, focusing on the noncooperative environment.

i. Dependence of social utility on past society behavior

It is natural in some social contexts to expect social utility to depend on the past level of the mean choice level. Examples of this feature would include intergenerational models of social norms in which offspring attitudes depend on the behavior of adult role models. A general formulation of this idea may be done using the analysis of section 3.*ii*, after incorporating the additional feature that the social utility parameter J depends on the lagged expected average choice level. In an equilibrium, the expected average choice level must solve

$$m_t = tanh(\beta h + \beta J(m_{t-1})m_{t-1}) \tag{34}$$

under either social utility specification. Fixed points of this equation will represent self-consistent steady states.

This equation, will, depending on the specification of $J(m_{t-1})$, be capable of exhibiting much more complicated behavior than the baseline model. For example, if $J(0) < \beta^{-1}$, whereas $J(m) > \beta^{-1}$ for $|m| > K$, then the model can (depending on K and h) exhibit a stable steady state at a mean level near 0 as well as at stable equilibria at mean levels near -1 and 1, unlike the analysis in section 3.*ii*. Additional unstable steady states can emerge as well.

ii. Asymmetric social utility

An alternative generalization of the social utility term would allow for an asymmetry in the consequence of a choice above the mean level of others versus a choice below this level. This would mean replacing J in (4) and (5) with J_+ if $\omega_i = 1$, J_- otherwise. Self-consistency would, under the noncooperative equilibrium with proportional spillovers, require that the mean choice level equals a root of

$$m^* = \frac{exp(\beta h + \beta J_+ m^*) - exp(-\beta h - \beta J_- m^*)}{exp(\beta h + \beta J_+ m^*) + exp(-\beta h - \beta J_- m^*)}. \tag{35}$$

Under a conformity effect, the self-consistent mean is a root of

$$m^* = \frac{exp(\beta h - \beta\frac{J_+}{2}(1 - m^*)^2) - exp(-\beta h - \beta\frac{J_-}{2}(-1 - m^*)^2)}{exp(\beta h - \beta\frac{J_+}{2}(1 - m^*)^2) + exp(-\beta h - \beta\frac{J_-}{2}(-1 - m^*)^2)}. \tag{36}$$

Hence the two solutions no longer coincide.

One interesting feature of these equations is that they illustrate how the relationship between large social utility effects in one direction and multiplicity of mean choice levels will depend critically on overall social utility specification. Suppose that $J_- = 0$, and consider the limiting behavior of (35) and (36) as $J_+ \Rightarrow \infty$. In the case of proportional spillovers, -1 and 1 are roots in the limit, whereas under conformity, -1 is still a root whereas 1 is not. Intuitively, while a large J_+ makes the choice of 1 under proportional spillovers extremely desirable for any positive mean, no such effect occurs under the conformity specification.

iii. Heterogeneity in deterministic private utility

A final extension would allow the $u(\cdot)$ term to vary across individuals. From the perspective of the development of the noncooperative equilibrium, this is equivalent to replacing the common h with different h_i's across individuals. Such heterogeneity will naturally arise when considering the econometric implementation of a model of this type, as will be seen below. We associate the empirical probability measure $dF_{h,I}(\cdot)$ with these individual characteristics. Reworking eqs. (8) to (12), it is straightforward to verify that a self-consistent mean for the noncooperative equilibrium must solve

$$m^* = \int tanh(\beta h + \beta J m^*) dF_{h,I}(h) \tag{37}$$

Existence of an equilibrium choice level will follow from the same argument given for Proposition 1 above; similarly, the equivalence of the proportional spillovers and conformity cases (Proposition 7) is unaffected. So long as $dF_{h,I}(\cdot)$ converges weakly to some probability measure $dF_h(\cdot)$, Proposition 6 can be generalized accordingly using techniques developed in Amaro de Matos and Perez (1991). Analysis of the other propositions will require the imposition of some restrictions on $dF_{h,I}(\cdot)$. For example, if $dF_{h,I}(\cdot)$ is symmetric, Brock and Durlauf (2000) show that $\beta J > 1$ is still necessary for multiple self-consistent solutions of m^*.

6. Econometrics

The model we have developed is capable of direct econometric implementation using standard methods.[8] Hence, the model may be subjected to hypothesis testing, which can reveal the importance of social interactions, as well as specification testing to see whether the basic framework is consistent with a given data set. In this discussion, we will ignore issues of asymptotics and focus on the identification of social interaction effects. More extensive discussion may be found in Brock and Durlauf (2000). For our binary choice model, we consider the identification based on a naive estimator of the parameters of the model. By naive, we refer to the case where a logistic regression is computed which does not impose the relationships between neighborhood means. In this case, the conditional likelihood function for the set of individual choices will have a standard logistic form.

We assume that each individual is drawn randomly from a set of neighborhoods. Within each neighborhood, all interactions are global. For notational purposes, we denote individuals as i and the neighborhood (which means the set of other individuals who influence i through interactions) as $n(i)$.

It is natural for empirical implementation that we relax the assumption that there exists a constant h which characterizes the private deterministic utility difference between the two choices. Instead, we assume that there exists an r-length vector of individual-specific observables $\underset{\sim}{X}_i$ and an s-length vector of exogenously determined neighborhood observables $\underset{\sim}{Y}_{n(i)}$ associated with each individual in the sample. This will allow us to replace the private utility component h with a general term h_i parametrized as

[8]Other empirical analyses have attempted to identify the presence of interactions based on spatial clumping in behavioral data. See Topa (1999) and Conley and Topa (1999) for analyses of this type. Alternatively, one can exploit the excess volatility of cross region behavior to identify social interactions, as done by Glaeser, Sacerdote and Scheinkman (1996). The advantage of our estimation approach, of course, is that it is structural and so all parameters have behavioral interpretations. Hence one can, in principle, compute the effects of policy changes. One advantage of the clumping/spatial correlation approach is that it can be applied to aggregate data, whereas our analysis requires individual level observations. Another possible advantage of this alternative relative to ours is robustness to misspecification. This is a topic which warrants further exploration.

$$h_i = k + c'\underset{\sim}{X}_i + d'\underset{\sim}{Y}_{n(i)}. \tag{38}$$

Notice that this specification means that none of the individual-specific observables $\underset{\sim}{X}_i$ or neighborhood observables $\underset{\sim}{Y}_{n(i)}$ contains a constant term. The difference between elements of $\underset{\sim}{X}_i$ versus $\underset{\sim}{Y}_{n(i)}$ is that any two agents in the same neighborhood must possess the same $\underset{\sim}{Y}_{n(i)}$'s although their $\underset{\sim}{X}_i$'s may differ.

Using our theoretical model of global interactions (and exploiting symmetry of the logistic density function), the likelihood is

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

$$\prod_i Prob(\omega_i = 1 \mid \underset{\sim}{X}_i, \underset{\sim}{Y}_{n(i)}, m^e_{n(i)})^{\frac{1+\omega_i}{2}} \cdot Prob(\omega_i = -1 \mid \underset{\sim}{X}_i, \underset{\sim}{Y}_{n(i)}, m^e_{n(i)})^{\frac{1-\omega_i}{2}} \sim$$

$$\prod_i (exp(\beta k + \beta c'\underset{\sim}{X}_i + \beta d'\underset{\sim}{Y}_{n(i)} + \beta J m^e_{n(i)})^{\frac{1+\omega_i}{2}}$$

$$\cdot\, exp(-\beta k - \beta c'\underset{\sim}{X}_i - \beta d'\underset{\sim}{Y}_{n(i)} - \beta J m^e_{n(i)})^{\frac{1-\omega_i}{2}}) \tag{39}$$

As is standard for logistic models, the complete set of model parameters is not identified as k, c', d' and J are each multiplied by β. We therefore proceed under the normalization $\beta = 1$.

The reason that .identification is a concern in a model like this is the presence of the term $m^e_{n(i)}$ in the likelihood function. Since this terms embodies a rationality condition, it is a function of other variables in the likelihood function. Specifically, we assume that

$$m^e_{n(i)} = m_{n(i)} = \int tanh(k + c'\underset{\sim}{X} + d'\underset{\sim}{Y}_{n(i)} + Jm_{n(i)})dF_{\underset{\sim}{X} \mid \underset{\sim}{Y}_{n(i)}}. \tag{40}$$

Here $F_{\underset{\sim}{X} \mid \underset{\sim}{Y}_{n(i)}}$ denotes the conditional distribution of $\underset{\sim}{X}$ in neighborhood $n(i)$

given the neighborhood characteristics $\underset{\sim}{Y}_{n(i)}$. What this means is that each agent is assumed to form the conditional probabilities of the individual characteristics in a neighborhood given the aggregates which determine his or her payoffs. Since one can always add elements of $\underset{\sim}{Y}_{n(i)}$ with zero coefficients to the payoff equation for agents, this is without loss of generality.

Rather than prove identification for the particular case where the theoretical model is logistic (see McFadden (1974) and Amemiya (1993, chapter 9) for proofs for this case) we prove identification for an arbitrary known distribution function for the random payoff terms. Specifically, we assume that the conditional probability of individual i's choice can be written as

$$Prob(\epsilon(\omega_i) - \epsilon(-\omega_i) \leq z \mid \underset{\sim}{X}_i, \underset{\sim}{Y}_{n(i)}, m^e_{n(i)}) =$$

$$F(z \mid k + c'\underset{\sim}{X}_i + d'\underset{\sim}{Y}_{n(i)} + Jm^e_{n(i)}) \tag{41}$$

where F is a known probability distribution function that is continuous and strictly increasing in z.

We consider identification based on a naive estimator of the parameters of the model. By naive, we refer to the situation where parameter estimates for the model are computed which do not impose the rational expectations condition between neighborhood means and neighborhood characteristics, but rather uses these variables as regressors. Hence, we assume that $m^e_{n(i)}$ is known to the researcher; see discussion below for the case when $m^e_{n(i)}$ is not observable.

To formally characterize identification, we employ the following notation. Define $supp(\underset{\sim}{X}, \underset{\sim}{Y}, m^e)$ as the joint support of the distribution of $(\underset{\sim}{X}_i, \underset{\sim}{Y}_{n(i)}, m^e_{n(i)})$. Intuitively, the definition of identification we employ says that a model is identified if there do not exist two distinct sets of parameter values each of which produces (for all subsets of $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ which occur with positive probability) identical probabilities for individual choices and which are also self-consistent.

Definition: Global identification in the binary choice model with interactions and self-consistent expectations

The binary choice model is globally identified if for all parameter pairs (k,c,d,J) and $(\bar{k},\bar{c},\bar{d},\bar{J})$

$$k + c'\underset{\sim}{X}_i + d'\underset{\sim}{Y}_{n(i)} + Jm^e_{n(i)} = \bar{k} + \bar{c}'\underset{\sim}{X}_i + \bar{d}'\underset{\sim}{Y}_{n(i)} + \bar{J}m^e_i \tag{42}$$

and

$$m^e_{n(i)} = m_{n(i)} =$$

$$\int \omega_i dF(\omega_i \mid k + c'\underset{\sim}{X} + d'\underset{\sim}{Y}_{n(i)} + Jm_{n(i)})dF_{\underset{\sim}{X} \mid \underset{\sim}{Y}_{n(i)}} =$$

$$\int \omega_i dF(\omega_i \mid \bar{k} + \bar{c}'\underset{\sim}{X} + \bar{d}'\underset{\sim}{Y}_{n(i)} + \bar{J}\, m_{n(i)})dF_{\underset{\sim}{X} \mid \underset{\sim}{Y}_{n(i)}} \tag{43}$$

$\forall\ (\underset{\sim}{X}_i,\underset{\sim}{Y}_{n(i)},m^e_{n(i)}) \in supp(\underset{\sim}{X},\underset{\sim}{Y},m^e)$ imply that $(k,c,d,J) = (\bar{k},\bar{c},\bar{d},\bar{J})$

In order to establish conditions under which identification can hold we follow the argument in Manski (1988), Proposition 5, and state the following Proposition, whose proof appears in the Technical Appendix. The assumptions we make are clearly sufficient rather than necessary; weakening the assumptions is left to future work. In interpreting the assumptions, note that Assumption i is the one used by Manski to identify this model when there are no endogenous effects, i.e. if J is known a priori to be 0. The assumption, of course does nothing more than ensure that the individual and contextual regressors are not linearly dependent. The additional assumptions are employed to account for the fact that $m_{n(i)}$ is a nonlinear function of the contextual effects.

Proposition 12. Sufficient conditions for identification to hold in the binary choice

model with interactions and self-consistent beliefs

Assume

i. $supp(\underset{\sim}{X}_i, \underset{\sim}{Y}_{n(i)})$ is not contained in a proper linear subspace of R^{r+s}.[9]

ii. $supp(\underset{\sim}{Y}_{n(i)})$ is not contained in a proper linear subspace of R^s.

iii. No element of $\underset{\sim}{X}_i$ or $\underset{\sim}{Y}_{n(i)}$ is constant.

iv. There exists at least one neighborhood n_0 such that conditional on $\underset{\sim}{Y}_{n_0}$, $\underset{\sim}{X}_i$ is not contained in a proper linear subspace of R^r.

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

vi. $m_{n(i)}$ is not constant across all neighborhoods n.

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

This proposition on identification in the binary choice model reaches a different conclusion from Manski's (1993) analysis of identification in linear models with social interactions. The reason why the two cases differ is of interest in understanding the general identification problem for interactions.

Manski studied what he referred to as the linear-in-means model, which maps into our notation as

$$\omega_i = c'\underset{\sim}{X}_i + d'\underset{\sim}{Y}_{n(i)} + Jm_{n(i)} + \epsilon_i. \tag{44}$$

[9]We follow Manski (1988) in defining a proper linear subspace of R^n as a space $X \subset R^n$ such that there is a vector $a \in R^n$ such that $X = a + L$ where L is a linear vector subspace of R^n with the dimension of L is less than n and where the notation $a + L$ denotes the set of all x's such that $x = a + l$ for some $l \in L$.

The unique self-consistent solution $m_{n(i)}$ for the linear-in-means model is easily computed by applying an expectations operator to both sides of the individual behavioral equation,

$$m_{n(i)} = \frac{c'E(\underset{\sim}{X}_i \mid \underset{\sim}{Y}_{n(i)}) + d'\underset{\sim}{Y}_{n(i)}}{1-J}. \tag{45}$$

(Note that we have preserved the information assumption used in eq. (40). Brock and Durlauf (2000) show to analyze identification under alternative information assumptions.) Hence for the linear-in-means model, $m_{n(i)}$ is a linear combination of various neighborhood-level variables. Manski studied the specific case where in the behavioral relation the analogous neighborhood level variable is always included for each individual-level variable, so that, for example, when one controls for individual education, one also controls for average neighborhood education. In this case, the linear space spanned by $E(\underset{\sim}{X}_i \mid \underset{\sim}{Y}_{n(i)})$ is the same linear space as that spanned by $\underset{\sim}{Y}_{n(i)}$, so $m_{n(i)}$ is linearly dependent on $\underset{\sim}{Y}_{n(i)}$, and the model is not identified.[10,11] On the other hand, as implied by Proposition 12, identification does hold for the binary choice model under the same assumption on the relationship between the individual and neighborhood-level contextual effects which produced nonidentification in the linear-in-means case.

Why is there this difference between the binary choice and the linear-in-means frameworks? The answer is that the binary choice framework imposes a nonlinear relationship between group characteristics and group behaviors whereas the linear-in-means model (by definition, of course) does not. Intuitively, suppose that one moves an individual from one neighborhood to another and observes the differences in his behavior. If the characteristics and behaviors of the neighborhoods always move in proportion as one moves across neighborhoods, then clearly one could not determine the respective roles of the characteristics as

[10]Brock and Durlauf (2000) show that a necessary condition for identification in the linear-in-means model is that $E(\underset{\sim}{X}_i \mid \underset{\sim}{Y}_{n(i)})$ is not contained in the linear space spanned by $\underset{\sim}{Y}_{n(i)}$.

opposed to the behavior of the group in determining individual outcomes. This can never happen in the logistic binary choice case given that the expected average choice must be bounded between -1 and 1. So, for example, if one moves across a sequence of richer and richer communities, the percentage of high school graduates cannot always increase proportionately with income.

Put differently, identification for binary choice models such as the logistic which transform a linear combination of some regressors into probabilities, require that the regressors used in computing probabilities not be linearly dependent, just as linear independence of regressors is required for identification in linear regressions. The relationship between the expected average neighborhood choice in a binary choice model and the regressors which characterize the causal determinants of individual behavior is necessarily nonlinear for sufficient variation in the neighborhood characteristics, given that probabilities are bounded between 0 and 1. The unboundedness condition *iii.* in Proposition 12 does precisely this by making sure that the variation in neighborhood characteristics is sufficient for any

[11]Following discussion in Brock and Durlauf (2000), we observe that the conditions for identification in the linear-n-means model bear a close relationship to the conditions for identification in rational expectations models as studied in Wallis (1980). This is apparent when one modifies the linear-in-means model so that it now describes behaviors at different points in time, i.e.

$$\omega_{i,t} = c'\underset{\sim}{X}_{i,t} + d'\underset{\sim}{Y}_{n(i),t} + Jm^e_{n(i),t} + \epsilon_{i,t}.$$

Let $\underset{\sim}{\omega}_t$ denote the column vector of choices at t, X_t and Y_t denote matrices whose columns are the $\underset{\sim}{X}_{i,t}$'s and $\underset{\sim}{Y}_{n(i),t}$'s respectively, and C and D denote conformable matrices whose rows are c' and d' respectively. Then a panel of observations on individuals can be written as

$$\underset{\sim}{\omega}_t = CX_t + DY_t + J\underset{\sim}{m}^e_t + \underset{\sim}{\epsilon}_t.$$

When $D = 0$, one has the vector linear-in-means model version of equation (2.1) in Wallis (1980). The particular identification problems associated with the linear-in-means model occur because of the need to identify D. The identification problem is particularly serious when the Y matrix consists of neighborhood averages of X_t, which is the insight of Manski (1993). We thanks James Heckman for alerting us to the importance of this relationship.

model relative to its alternative. Thus the regressor $m_{n(i)}$ in the binary choice model cannot depend linearly on the other regressors in the way it does in the linear-in-means model.

Finally, we note two limitations in this analysis. First, our argument has proceeded under the assumption that there is no endogenous self-selection into neighborhoods. Accounting for such endogeneity of memberships does not imply that identification is impossible; rather, it means that one will need to have available suitable instruments to control for endogeneity (see Evans, Oates, and Schwab (1992)) or an explicit self-selection correction of the type pioneered by Heckman (1979). As described in Brock and Durlauf (2000), the nonlinearity argument we have developed for the binary choice model also applies to such environments.

Second, we have proceeded under the assumption that data are available for the neighborhoods which actually define the relevant social interactions environment. When these groups are not known, another dimension of identification needs to be addressed. Manski (1993) notes the difficulties which adhere in empirical work in which neighborhood structures need to be inferred along with the strength of interactions within neighborhoods. Whether any inferences on group effects and group interaction structure can be made without prior information on neighborhood structure is not known. At a minimum, it seems clear that efforts to generate data from which to learn about interaction effects should attempt to identify respondents' perspectives on which groups matter.

However, we do note that there are circumstances in which one may be interested in the effects of a given interaction structure on outcomes, such as the effect of school district composition on students, in which one has a set of groups whose compositions are subject to policy interventions, so that the goal of the analysis is to predict how these particular groups affect outcomes. In such cases, the analysis we describe will apply, subject to accounting for the effects of using groups which only approximate actual interaction environments on the analysis.

7. Summary and conclusions

This paper has developed a framework for characterizing discrete decisions when individuals experience private as well as social utility from their choices. The model is shown to produce a number of interesting features. First, multiple, locally stable equilibrium levels of average behavior are shown to exist when social utility effects are large enough and decisionmaking is noncooperative. Second, a large social multiplier can exist in terms of relating small changes in private utility to large equilibrium changes in average behavior. Third, while the social planner eliminates the multiplicity of average outcomes, other features of the noncooperative equilibrium, such as the presence of a large social multiplier, are preserved. Fourth, the model provides some insights into a number of empirical phenomena. Fifth, the model is econometrically tractable as its equilibrium is mathematically equivalent to a logistic likelihood function. While the presence of multiple equilibria and social multipliers are common features of models with social interactions, our ability to introduce heterogeneity and uncertainty into the microeconomic specification of decisionmaking and the direct link between the theoretical model and an econometrically-implementable likelihood function that is thereby induced are, we believe, unique to this literature.

In terms of future research, several areas of investigation seem especially important. First, there needs to be further study of models with self-selection. The analysis of this paper has taken the interactions group as given and then explored the properties of behavior within the group. A natural extension would consider the theoretical and econometric consequences of analyzing environments in which groups are endogenously determined. In particular, it would be valuable to integrate the social utility analysis of the current paper with a framework such as Bénabou (1993, 1996) or Durlauf (1996a,b), which allows for endogenous selection of one's reference group. This integration would enhance the ability of the current framework to explain phenomena such as the emergence and perpetuation of ghettos. Second, it is important to extend the binary choice analysis to a longitudinal framework. Behaviors such as dropping out of school or illegal activity may be binary at a point in time, but are best conceptualized using survival analysis or other longitudinal methods. Some results on the econometric identification of these types of models are found in Brock and Durlauf (2000), but much more remains to be done both in terms of econometrics as well as theory.

Technical Appendix

1. Proof of Proposition 8

The asymptotic properties of the probability measure described by eq. (25) can be analysed by using the following identity, whose usefulness was exploited by Kac (1968),

$$exp(a^2) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} exp(-\frac{x^2}{2} + 2^{1/2} x a) dx. \qquad \text{(A.1)}$$

This identity can be verified immediately by dividing both sides of the expression by $exp(a^2)$ and observing that eq. (A.1) is equivalent to the statement that the integral of the probability density of a normal $(2^{1/2}a, 1)$ random variable over its support is 1. When this identity is employed, substituting a with $(\frac{\beta J}{I})^{1/2} \sum_{i=1}^{I} \omega_i$, into (A.1), we have

$$exp((\frac{\beta J}{I})^{1/2} \sum_{i=1}^{I} \omega_i)^2 = (2\pi)^{-1/2} \int_{-\infty}^{\infty} exp(-\frac{x^2}{2} + x(\frac{2\beta J}{I})^{1/2} \sum_{i=1}^{I} \omega_i) dx. \quad \text{(A.2)}$$

Using the change of variable $y = (2\beta J/I)^{1/2} x$, it is therefore the case that

$$exp(\beta(\sum_{i=1}^{I} h\omega_i + \frac{J}{I}(\sum_{i=1}^{I} \omega_i)^2)) =$$

$$exp(\beta(\sum_{i=1}^{I} h\omega_i) \cdot (\frac{I}{4\pi\beta J})^{1/2} \int_{-\infty}^{\infty} exp(-\frac{Iy^2}{4\beta J} + y \sum_{i=1}^{I} \omega_i) dy) =$$

$$\prod_{i=1}^{I} exp(\beta h \omega_i) \cdot (\frac{I}{4\pi\beta J})^{1/2} \int_{-\infty}^{\infty} exp(-\frac{Iy^2}{4\beta J}) \prod_{i=1}^{I} exp(y\omega_i) dy =$$

$$(\frac{I}{4\pi\beta J})^{1/2} \int_{-\infty}^{\infty} exp(-\frac{Iy^2}{4\beta J}) \prod_{i=1}^{I} exp((\beta h + y)\omega_i) dy. \qquad \text{(A.3)}$$

Summing this expression over all possible realizations of $\underset{\sim}{\omega}$ yields

$$\sum_{\nu_1 \in \{-1,1\}} \cdots \sum_{\nu_I \in \{-1,1\}} exp(\beta(\sum_{i=1}^{I} h\nu_i + \frac{J}{I}(\sum_{i=1}^{I} \nu_i)^2)) =$$

$$\sum_{\nu_1 \in \{-1,1\}} \cdots \sum_{\nu_I \in \{-1,1\}} (\frac{I}{4\pi\beta J})^{1/2} \int_{-\infty}^{\infty} exp(-\frac{Iy^2}{4\beta J}) \prod_{i=1}^{I} exp((\beta h + y)\nu_i) dy =$$

$$(\frac{I}{4\pi\beta J})^{1/2} \int_{-\infty}^{\infty} exp(-\frac{Iy^2}{4\beta J}) \cdot$$

$$(\sum_{\nu_1 \in \{-1,1\}} exp((\beta h + y)\nu_1)) \cdot \ldots \cdot (\sum_{\nu_I \in \{-1,1\}} exp((\beta h + y)\nu_I)) dy. \quad \text{(A.4)}$$

However, since

$$\sum_{\nu_i \in \{-1,1\}} exp((\beta h + y)\nu_i) = exp(\beta h + y) + exp(-\beta h - y) \ \forall \ i, \quad \text{(A.5)}$$

$Prob(\underset{\sim}{\omega})$ will equal

$$\frac{\int_{-\infty}^{\infty} exp(-\frac{Iy^2}{4\beta J}) \prod_{i=1}^{I} exp((\beta h + y)\omega_i) dy}{\int_{-\infty}^{\infty} (exp(-\frac{y^2}{4\beta J})(exp(\beta h + y) + exp(-\beta h - y)))^{I} dy} \quad \text{(A.6)}$$

which can be rewritten as

$$Prob(\underset{\sim}{\omega}) = \int_{-\infty}^{\infty} K(I,y) \frac{exp((\beta h + y)(\sum_{i=1}^{I} \omega_i))}{(exp(\beta h + y) + exp(-\beta h - y))^{I}} dy \quad \text{(A.7)}$$

where

$$K(I,y) = \frac{(exp(-\frac{y^2}{4\beta J})(exp(\beta h + y) + exp(-\beta h - y)))^{I}}{\int_{-\infty}^{\infty} (exp(-\frac{y'^2}{4\beta J})(exp(\beta h + y') + exp(-\beta h - y')))^{I} dy'}. \quad \text{(A.8)}$$

Consider the function $K(I,y)$. Clearly, $\int_{-\infty}^{\infty} K(I,y)dy$ will equal 1 for all I. Further, the shape of the function with respect to y for fixed I will be determined by

$$(exp(-\frac{y^2}{4\beta J})(exp(\beta h + y) + exp(-\beta h - y)))^I \qquad (A.9)$$

since the denominator of (A.8) is independent of y. As I increases, the ratio of the value of $K(I,y)$ evaluated at y^*, defined by

$$y^* = max_y(exp(-\frac{y^2}{4\beta J})(exp(\beta h + y) + exp(-\beta h - y))), \qquad (A.10)$$

to the value of $K(I,y)$ at any other y should become arbitrarily large, so long as y^* is unique, since we are taking a function to the I'th power. Making this rigorous is a straightforward exercise using the following result, found in Murray (1984), pg. 34.

Approximation Theorem.

Let $H(t)$ be a function on the interval (a,b) which takes a global maximum at a point α in the interval and let $H(t)$ be smooth enough to possess a second-order Taylor expansion at point α with $H''(\alpha) < 0$. Let $G(t)$ denote a continuous function. Then

$$\int_{-\infty}^{\infty} G(t)exp(IH(t))dt =$$

$$exp(IH(\alpha))(G(\alpha)(\frac{-2\pi}{IH''(\alpha)})^{1/2}) + O(I^{-3/2}) \qquad (A.11)$$

This formula states, in a precise way, the sense in which the mass of the integral piles up at the maximizer α as $I \Rightarrow \infty$. Differentiating and rearranging terms of eq. (A.9) therefore implies that y^* must be a root of

$$y^* = 2\beta J tanh(\beta h + y^*). \quad \text{(A.12)}$$

Since y^* is a global maximum, the root of the first-order condition eq. (A.12) which also solves eq. (A.10) must be the one that has the same sign as h, so that uniqueness is assured so long as $h \neq 0$. We will assume that h is nonzero for the subsequent analysis.

Intuitively, this discussion leads one to expect $K(I,y)$ to asymptotically behave as a Dirac function. Further, given the term

$$\frac{exp((\beta h + y)(\sum_{i=1}^{I} \omega_i))}{((exp(\beta h + y) + exp(-\beta h - y)))^I} = \prod_{i=1}^{I} \frac{exp((\beta h + y)\omega_i)}{exp(\beta h + y) + exp(-\beta h - y)} \quad \text{(A.13)}$$

in (A.7), one would expect that $K(I,y)$ acts on this term in such a way that the probability measure for $\underset{\sim}{\omega}$ will possess the property that

$$\bar{\omega}_I \Rightarrow_w \frac{exp(\beta h + y^*) - exp(-\beta h - y^*)}{exp(\beta h + y^*) + exp(-\beta h - y^*)} = tanh(\beta h + y^*). \quad \text{(A.14)}$$

This heuristic argument can be formalized using LaPlace's method, (see Erdélyi (1956), section 2.4 for a general exposition and Kac (1968) for the development of the method in the context of the Curie-Weiss model). The actual application of the method in the current context is in fact quite subtle and was originally analyzed in Brock (1993), pg. 22. Combining eqs. (A.10) and (A.12), and rewriting y^* as $2\beta J m^*$ leads to Proposition 8.

Finally, there is an interesting connection between our solution to the behavior of a social planner and the maximization of social surplus as analyzed in McFadden (1981). Following McFadden, social surplus will equal $\sum_i (u(\omega_i) - \frac{J}{2}(\omega_i - \bar{\omega}_I)^2)$. Following (25), the probability measure of the social surplus can be expressed as a function of $G(\underset{\sim}{\omega}) = \sum_i h\omega_i + \frac{J}{2I}(\sum_i \omega_i)^2$. Then it can be shown (Brock (1993)) that

$$\beta(lim_{I\Rightarrow\infty}E(max_{\underset{\sim}{\omega}}\, I^{-1}G(\underset{\sim}{\omega}))) = lim_{I\Rightarrow\infty}(I^{-1}lnZ_I) =$$

$$max_y\; ln(exp(-\frac{y^2}{2\beta J})M(\beta h + y)) =$$

$$max_m\; ln(exp(-\frac{(\beta Jm)^2}{2\beta J})M(\beta h + \beta Jm)) \tag{A.15}$$

where in this statement

$$Z_I = \sum_{\nu_1\in\{-1,1\}} \cdots \sum_{\nu_I\in\{-1,1\}} exp(\beta(\sum_{i=1}^{I} h\nu_i + \frac{J}{2I}(\sum_{i=1}^{I}\nu_i)^2)) \tag{A.16}$$

and

$$M(s) = exp(s) + exp(-s). \tag{A.17}$$

As would be expected, one maximizes a notion of social welfare in the large economy limit in order to find the socially optimal states.

2. Proof of Proposition 12

For a given parameter set (k,c,d,J), assume by way of contradiction that there exists an alternative $(\bar{k},\bar{c},\bar{d},\bar{J})$ such that on $supp(\underset{\sim}{X},\underset{\sim}{Y},m^e)$ we have

$$(k-\bar{k}) + (c'-\bar{c}')\underset{\sim}{X}_i + (d'-\bar{d}')\underset{\sim}{Y}_{n(i)} + (J-\bar{J})m^e_{n(i)} = 0 \tag{A.18}$$

and

$$m^e_{n(i)} = m_{n(i)} = \int \omega_i dF(\omega_i \,|\, k + c'\underset{\sim}{X} + d'\underset{\sim}{Y}_{n(i)} + Jm_{n(i)})dF_{\underset{\sim}{X}\,|\,\underset{\sim}{Y}_{n(i)}} =$$

$$\int \omega_i dF(\omega_i \,|\, \bar{k} + \bar{c}'\underset{\sim}{X} + \bar{d}'\underset{\sim}{Y}_{n(i)} + \bar{J}m_{n(i)})dF_{\underset{\sim}{X}\,|\,\underset{\sim}{Y}_{n(i)}}. \tag{A.19}$$

Notice the Proposition is true if it is the case that $J - \bar{J}$ is zero. Otherwise, $\underset{\sim}{X}_i$ and $\underset{\sim}{Y}_{n(i)}$ would lie in a proper linear subspace of R^{r+s} which violates Assumption *i*. To show that the parameter vector c is identified, notice that equation (A.18) implies that for elements of $supp(\underset{\sim}{X}, \underset{\sim}{Y}, m^e)$, conditional on $\underset{\sim}{Y}_{n(i)}$

$$(c' - \bar{c}')\underset{\sim}{X}_i = \rho(\underset{\sim}{Y}_{n(i)}) \tag{A.20}$$

where $\rho(\underset{\sim}{Y}_{n(i)}) = -(k - \bar{k}) - (d' - \bar{d}')\underset{\sim}{Y}_{n(i)} - (J - \bar{J})m^e_{n(i)}$. (A.20) must hold for all neighborhoods, including n_0 as described in Assumption *iv.* of the Theorem. This would mean that, conditional on $\underset{\sim}{Y}_{n_0}$, and given that $\underset{\sim}{X}_i$ cannot contain a constant by Assumption *iii.*, that $\underset{\sim}{X}_i$ is contained in a proper linear subspace of R^r and therefore violates the Assumption *iv.* of the Proposition. Hence, c is identified.

Given identification of c, (A.18) now implies, if $J \neq \bar{J}$, that $m^e_{n(i)}$ is a nontrivial linear function of $\underset{\sim}{Y}_{n(i)}$, unless $(d' - \bar{d}')$ and/or $m^e_{n(i)}$ is always equal to zero. The latter is ruled out by Assumption *vi*. Linear dependence of $m^e_{n(i)}$ on $\underset{\sim}{Y}_{n(i)}$ when $(d' - \bar{d}') \neq 0$ contradicts the combination of the requirement that support of $m^e_{n(i)}$ is $[-1,1]$ with Assumption *v.*, that the support of each component of $\underset{\sim}{Y}_{n(i)}$ is unbounded, since $\underset{\sim}{Y}_{n(i)}$ can, if it is unbounded, assume values with positive probability that violate the bounds on $m^e_{n(i)}$. So, J is identified. If J is identified and $(d' - \bar{d}') \neq 0$, then (A.18) requires that

$$(d' - \bar{d}')\underset{\sim}{Y}_{n(i)} = -(k - \bar{k}) \tag{A.21}$$

for all $\underset{\sim}{Y}_{n(i)} \in supp(\underset{\sim}{Y}_{n(i)})$. This implies, since by Assumption *iii.* $\underset{\sim}{Y}_{n(i)}$ does not contain a constant, that $supp(\underset{\sim}{Y}_{n(i)})$ is contained in a proper linear subspace of R^s, which contradicts condition *ii.* of the Proposition. Therefore, $d' = \bar{d}'$. This immediately implies that $k = \bar{k}$ and the Proposition is verified.

Bibliography

Amaro de Matos, J. M. G., and J. F. Perez, (1991), "Fluctuations in the Curie-Weiss Version of the Random Field Ising Model," Journal of Statistical Physics, 62, 3/4, 587-608.

Amemiya, T., (1985), Advanced Econometrics, Cambridge: Harvard University Press.

Anderson, E., (1990), Streetwise, Chicago: University of Chicago Press.

Anderson, S., A. de Palma, and J.-F. Thisse, (1992), Discrete Choice Theory of Product Differentiation, Cambridge: MIT Press.

Arrow, K. and F. Hahn, (1971), General Competitive Analysis, New York: Holden Day.

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.

Arthur, W. B., (1989), "Increasing Returns, Competing Technologies and Lock-In by Historical Small Events: The Dynamics of Allocation Under Increasing Returns to Scale," Economic Journal, 99, 116-131.

Becker, G., (1991), A Treatise on the Family: Enlarged Edition, Cambridge: Harvard University Press.

Becker, G., (1996), Accounting for Tastes, Cambridge: Harvard University Press.

Bell, A., (1994), "Dynamically Interdependent Preferences in a General Equilibrium Framework," mimeo, University of Wisconsin at Madison.

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

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

Bernheim, D., (1994), "A Theory of Conformity," Journal of Political Economy, 5, 102, 841-877.

Blume, L., (1993), "The Statistical Mechanics of Strategic Interaction," Games and Economic Behavior, 5, 387-424.

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

Bowles, S., and H. Gintis, (1994), "Efficient Redistribution: New Rules for Markets, States, and Communities," mimeo, University of Massachusetts at Amherst.

Brock, W., (1993), "Pathways to Randomness in the Economy: Emergent Nonlinearity and Chaos in Economics and Finance," Estudios Economicos, 8, 1, 3-55 and Social Systems Research Institute Reprint #410, Department of Economics, University of Wisconsin at Madison.

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

Brock, W., and C. Hommes, (1997), "Rational Routes to Randomness," Econometrica, 65, 1059-1096.

Bryant, J. (1983), "A Simple Rational Expectations Keynes-Type Model," Quarterly Journal of Economics, XCVIII, 525-529.

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

Coleman, J., (1988), "Social Capital in the Creation of Human Capital," American Journal of Sociology, 94 (supplement), S95-S120.

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

Conley, T. and G. Topa, (1999), "Socio-Economic Distance and Spatial Patterns in Unemployment," mimeo, New York University.

Cooper, R., and A. John, (1988), "Coordinating Coordination Failures in Keynesian Models," Quarterly Journal of Economics, CIII, 441-464.

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

Dasgupta, P., (1995), "The Population Problem: Theory and Evidence," Journal of Economic Literature, 33, 1879-1982.

Debreu, G., (1952), "A Social Equilibrium Existence Theorem," Proceedings of the National Academy of Sciences, 38, 886-893.

Durlauf, S., (1993), "Nonergodic Economic Growth," Review of Economic Studies, 60, 349-366.

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: Proceedings of the Ninth International Symposium on Economic Theory and Econometrics, W. Barnett, G. Gandolfo, and C. Hillinger, eds., Cambridge University Press.

Durlauf, S., (1997), "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.

Durlauf, S. and J. Walker, (1998), "Social Interactions and Fertility Transitions," mimeo, University of Wisconsin at Madison.

Ellis, R., (1985), Entropy, Large Deviations, and Statistical Mechanics, New York: Springer-Verlag.

Erdélyi, A., (1956), Asymptotic Expansions, New York: Dover Publications.

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

Föllmer, H., (1974), "Random Economies with Many Interacting Agents," Journal of Mathematical Economics, 1, 51-62.

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

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.

Ioannides, Yannis M., (1990), "Trading Uncertainty and Market Structure," International Economic Review, 31, 619-638.

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.

Kac, M., (1968), "Mathematical Mechanisms of Phase Transitions," in Statistical Physics: Phase Transitions and Superfluidity, vol. 1, M. Chretien, E. Gross, and S. Desar eds., Brandeis University Summer Institute in Theoretical Physics 1966.

Kohler, H.-P., (1997), Fertility and Social Interaction: An Economic Approach, PhD Dissertation, Department of Economics, University of California, Berkeley.

Krugman, P., (1995), The Self-Organizing Economy, unpublished manuscript, Stanford University.

Lewis, O., (1966), La Vida: A Puerto Rican Family in the Culture of Poverty, New York: Random House.

Liebow, E., (1967), Tally's Corner, Boston: Little, Brown.

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.

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

Manski, C., (1993a), "Identification Problems in the Social Sciences," in Sociological Methodology, vol. 23, P. Marsden, ed., Cambridge: Basil Blackwell.

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

Massey, D., and N. Denton, (1993), American Apartheid, Cambridge: Harvard University Press.

McFadden, D., (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, Frontiers in Econometrics, New York: Academic Press.

McFadden, D., (1984), "Econometric Analysis of Qualitative Response Models," in Z. Griliches and M. Intriligator, eds., Handbook of Econometrics: Volume II, Amsterdam: North-Holland.

Milgrom, P., and D. J. Roberts, (1990), "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," Econometrica, 58, 1255-1277.

Miyao, T., (1978a), "Dynamic Instability of a Mixed City in the Presence of Neighborhood Externalities," American Economic Review, 68, 454-463

Moffitt, R., (2000), "Policy Interventions, Low-Level Equilibria, and Social Interactions," mimeo, Johns Hopkins University.

Montgomery, J., (1994), "Revisiting *Tally's Corner*: Mainstream Norms, Cognitive Dissonance, and Underclass Behavior," Rationality and Society, 6, 462-488.

Murray, J., (1984), Asymptotic Analysis, New York: Springer-Verlag.

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.

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

Schelling, T., (1973), "Hockey Helmets, Concealed Weapons, and Daylight Saving: A Study of Binary Choices with Externalities," Journal of Conflict Resolution, 17, 381-428.

Spitzer, F., (1971), Random Fields and Interacting Particle Systems, Providence: Mathematical Association of America.

Steinberg, L. with B. Brown and S. Dornbusch, (1996), Beyond the Classroom, New York: Simon and Schuster.

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.

Weidlich, W., (1992), "Physics and Social Science: The Approach of Synergistics," Physics Reports, 204, 1-163.

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