QUADERNI

Università degli Studi di Siena
DIPARTIMENTO DI ECONOMIA POLITICA

GIULIO ZANELLA

Discrete Choice with Social Interactions and Endogenous Memberships

n. 442 - Novembre 2004

Abstract - This paper tackles the issue of self-selection in social interactions models. I develop a theory of sorting and behavior, when the latter is subject to social influences, extending the model developed by Brock and Durlauf (2001a, 2003) to allow for equilibrium group formation. Individuals choose a group, and a behavior subject to an endogenous social effect. The latter turns out to be a segregating force, and stable equilibria are stratified. The sorting process may induce, inefficiently, multiple behavioral equilibria. Such a theory serves as a means to solve identification and selection problems that may undermine the empirical detection of social effects on individual behavior. I exploit the theoretical model to build a nonlinear (in the social effect) selection correction term. Such a term allows identification, and solves the selection problem that arises when individuals can choose the group whose effect the researcher is trying to disentangle. The resulting econometric model, although relying on strict parametric assumptions, indicates a viable alternative when reliable instrumental variables are not available, or randomized experiments not possible.

JEL Classification Codes: C25, D85, E19, Z13, Z19

Keywords: social interactions, neighborhood effects, sorting, self-selection, nested logit, identification of social effects.

Acknowledgements. This paper was conceived and completed during my stay at the University of Wisconsin-Madison, Department of Economics, whose hospitality I acknowledge with gratefulness. I am greatly indebted to Ethan Cohen-Cole, Steven Durlauf, Yannis Ioannides, Nick Mader, Giacomo Rondina, and Adriaan Soetevent for invaluable suggestions and discussions.
I am also grateful to seminar participants at UWMadison and the EEA 2004 Madrid Congress for helpful comments. All remaining errors and nonsense are, of course, my own. This research is supported by a grant from the Program of Fellowships for Junior Scholars, MacArthur Research Network on Social Interactions and Economic Inequality.

Giulio Zanella, Department of Economics, Università di Siena, and University of Wisconsin-Madison.

To choose a neighborhood is to choose neighbors

-Thomas Schelling, *Dynamic Models of Segregation*, 1971[1]

1 Introduction

There is considerable interest in economics in models with social interactions or neighborhood effects, i.e. models that allow for interdependencies, not mediated by markets and enforceable contracts, between individual behavior and the behavior and characteristics of other individuals within neighborhoods, schools, or any other kind of reference group. This interest is mainly due to the fact that these models allow us to incorporate rigorously into economic theory concepts like social norms, social capital and peer effects, which are often regarded as loose concepts by economists, but which may have important aggregate effects[2]. This in turn allows us to study theoretically and empirically their economic effects addressing, respectively, endogeneity and identification problems within equilibrium frameworks. Reciprocal feedbacks between choice of behavior and choice of group are likely in presence of social interactions. For instance, if parents know the persons their kids associate with will somehow influence their behavior, they have a reason to carefully select the social environment their kids will mix with. Choosing a group means choosing social interactions: for this reason, considerations of selection are inescapable in models with social effects.

Yet, in the existing literature these two aspects of the problem are not fully integrated into a framework that is flexible enough for both theoretical and empirical analysis. A popular class of models has introduced the notion of social interactions in a Tiebout-type framework. These models consider group memberships as endogenous, in presence of social spillovers or complementarities (in the sense of Cooper and John, 1988, and Milgrom and Roberts, 1990), but either consider social interactions as exogenous in the Manski (1993) sense (e.g. de Bartolome, 1990, Bénabou, 1996, and Durlauf, 1996), or limit heterogeneity to an observable type (e.g. Bénabou, 1993, and Becker and Murphy, 2000). Other models offer an empirical framework for the analysis of group membership in the presence of social effects, but do not specify the structure of interactions at the micro level (e.g. Bayer and Tim-

[1] Quotation from Schelling (1971), p. 145.

[2] For an overview see Manski (2000) and Durlauf (2004).

mins, 2005). Another well-known class of models, well represented by Brock and Durlauf (2001a), uses a random utility framework to link theory and econometrics of social interactions, but treats memberships as exogenous.

This paper is an attempt to bridge these models and produce an integrated theory of social interactions and endogenous memberships that also suggests useful econometric specifications. I describe a model of choice of behavior subject to social interactions within groups, and choice between costly memberships in alternative groups. The model is simply an extension of Brock and Durlauf (2001a) to equilibrium group formation and of Benabou (1993) to a richer structure of unobservables, as well as of Benabou (1996) to endogenous social interactions. The goal of the paper is to illustrate the implications such integrated framework has for a complete theory of social interactions, as well as for the possibility to correctly identify social effects. On the theoretical side, I develop an idea first devised by Brock and Durlauf (2005) and use a special random utility model, the nested logit model whose structure can nicely accommodate the process of sequential and interdependent choice of group and behavior, to capture the essential interconnection between sorting and behavior in the presence of social interactions. The model implies that the sorting process and the social effect feedback reciprocally and interact in characterizing the equilibrium distribution of memberships and behaviors in the population. Positive spillovers and complementarities between individual behaviors turn out to be a stratifying force, a well known result in the endogenous neighborhood literature. Equilibrium stratification, along an individual observable trait, is induced by social interactions and is sustained by the equilibrium in the memberships market. The sorting process may generate an endogenous change in the qualitative properties of the system, with the emergence of "social traps". The decentralized outcome of sorting is generally inefficient.

This model synthesizes the aforementioned research programs in a very specific sense: it reduces to the Brock and Durlauf (2001a) model when individuals are not given a choice over memberships, and it reproduces the relevant results of Benabou (1993 and 1996) and Durlauf (1996), concerning the endogenous emergence of segregation and economic barriers between groups, when the stochastic component of group preferences is shut off. On the econometric side, such a synthesis indicates a way to solve in a microfounded way two major problems that typically arise in the empirical analysis of social effects, namely identification failure in linear models and selection bias. As suggested by Blume and Durlauf (2006), modelling selection within the statistical analysis is preferable to merely instrumenting for the variables that are indirectly chosen through sorting, which is a popular choice in the empirical social interactions literature. Another possibility to eliminate selection-bias is to use natural experiments (e.g. Sacerdote, 2001, and Zimmerman, 2003) or controlled experiments (e.g. Ichino and Falk, 2006), where people are exogenously assigned to groups. However, a proper experimental setting may not be available for particular applications, let alone extrapolations from existing experimental evidence. The advantage of the procedure I suggest is that it is derived from maximizing behavior, and equilibrium[3]. The main drawback is that it relies on very restrictive parametric assumptions. This is not much of a problem for theoretical purposes, but it may render the identification strategy of limited use for empirical applications. I will argue that such parametric solution can be of help in two directions. First, in performing a simple equilibrium-based selection test in applied work, which is something typically neglected in non-experimental empirical studies, and second, in thinking about semiparametric solutions to selection correction.

The paper is organized as follows. In section 2, I build the model in a very simple setting: two types of random-utility maximizing individuals, and a two-stage choice process: first one of two groups, on a competitive memberships market with fixed supply, then a binary behavior subject to social effects. In section 3, using the random utility maximization hypothesis, I derive a nested logit model of choice under self-consistent (i.e. rational) expectations. Section 4 analyzes the three components of an aggregate equilibrium: memberships market equilibrium, sorting equilibrium, and behavioral equilibrium. I establish existence, elucidate the relation between the three components of equilibrium, and discuss the conditions under which the sorting process induces multiple behavioral equilibria. Section 5 characterizes the sorting equilibrium. I show that the only stable equilibria are characterized by stratification along the individual type and that, in equilibrium, groups are more stratified when interactions are stronger, and when the unobservable determinants of sorting have low dispersion. Integration is an equilibrium, but an unstable one. Section 6 discusses efficiency issues. It is shown that integration may be the only efficient arrangement of individuals across groups when these are similar enough, and that, in general, the

[3] The procedure has been implemented by Ioannides and Zabel (2004) to estimate a model of social interactions in housing demand, but their approach is not based on equilibrium sorting.

decentralized equilibrium is inefficient. Finally, in section 7, I turn to the econometric issues, and show how the theoretical model can help construct econometric models to identify social interactions free of selection-bias, without using instrumental variables. Section 8 concludes. A few formal details and proofs are gathered in a technical appendix.

2 The model

Imagine a population of I individuals. Each individual is characterized by an observable real parameter $h \in \Theta$, referred to as the individual type. Individual i's type is denoted with h_i, $i = 1, ..., I$. The set of possible types is $\Theta = \{H, L\}$, with $H > L > 0$. Types are distributed in the population according to the discrete distribution f_h. Each individual is member of a group, indexed by $g \in G$, where $G = \{A, B\}$. The two groups have fixed capacities, I_A and I_B, such that $I_A + I_B = I$. Relative capacities are denoted $\alpha_A \equiv I_A/I$ and $\alpha_B \equiv I_B/I$. Groups are characterized by a set of exogenous variables, which we can think of as amenities, summarized by an index k_g. Each group charges a "membership fee", ρ_g, determined on a competitive memberships market with fixed supply. Fees end up outside the model. Denote with f_{hg} the distribution of h-types in group g. By definition, the distributions of types in the population and in any group g satisfy:

$$\begin{aligned} f_H + f_L &= 1 \\ f_{Hg} + f_{Lg} &= 1. \end{aligned} \tag{1}$$

Furthermore, the number of h-type individuals in group A, equal to $I_A f_{hA}$, plus the number of individuals of the same type in group B, $I_B f_{hB}$, must be equal to the number of h-type individuals in the population, $I f_h$. This implies that, for any type h:

$$\alpha_A f_{hA} + \alpha_B f_{hB} = f_h. \tag{2}$$

The individual problem is to choose a group $g \in \{A, B\}$, and a binary behavior $\omega \in \{-1, 1\}$. I assume there are interdependencies within groups:

when choosing behavior, individuals influence each other, in the sense to be specified below. Agents are assumed to be forward looking: when choosing membership they will take into account that group composition will affect the behavioral problem. Therefore choices over memberships are interdependent too, which is a key feature of the model. Choice can be represented as a two-stage process[4], as illustrated in figure 1. First a group is chosen and then behavior, given group composition. Choices are simultaneous across the population at each stage.

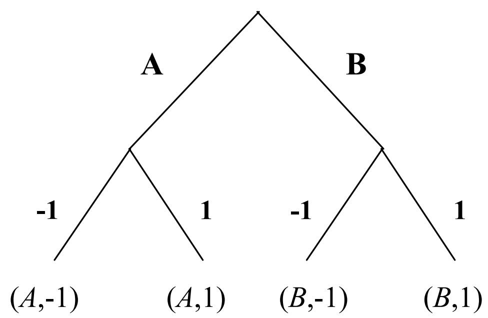

Figure 1. Individual decision tree.

In specifying preferences, I will extend the binary choice model with social interactions of Brock and Durlauf (2001a), to accommodate the choice structure depicted in figure 1. It is an extension in the sense that the choice set in their model is $\{-1, 1\}$, i.e. group composition is given, but is $\{A, B\} \times \{-1, 1\}$ here, i.e. group composition is endogenous. Therefore the model reduces to theirs when memberships are given. In models with social interactions, utility is usually separable in a private and social component:

$$V_i(g, \omega) = u(g, \omega, h_i) + s\left(\omega, m^e_{ig}\right) + \xi_{ig\omega} \tag{3}$$

In this case private utility has a deterministic type-specific part, $u(.)$, and an unobservable idiosyncratic component, ξ, and both depend on membership and behavior. Social utility, $s(.)$, depends on one's behavior and

[4]In what follows it doesn't make any difference whether group and behavior are chosen at once or sequentially. In many applications, the latter interpretation is more appropriate, which is why I stress the sequential character of choice.

expected mean behavior in the group, m^e_{ig}, which in turn depends on group composition[5]. Agent i observes $\xi_{ig\omega}$, but nobody else does. I also assume this random term is identically and independently distributed across individuals, and that something, to be specified in a moment, about this distribution is common knowledge.

It may be useful to refer to a couple examples at this point. Think of a household in which kids make a binary choice that influence their human capital (e.g. drop out of school, study hard, go to college) or that exposes them to risks (e.g. join a gang, commit crimes, use drugs, engage in risky sexual behavior), subject to such social effects as role models in the neighborhood or peer effects in schools. Parents can choose where to live or the school their kids will attend, i.e. have some control over kids' social interactions. In this case the individual type may be parents' human capital, private utility reflects kids' human capital or lifetime income as well as the amenities or prestige of living or sending kids in a specific place or school for parents, and the random term captures things like kids' talent and parents' attachment to specific places or schools, e.g. for cultural or ideological reasons[6]. Or think of an individual who chooses her or his social ties and then decides whether to work or live on welfare, or whether to behave honestly or infringe the law, when the social network is a source of job market information and welfare stigma, or of sense of lawfulness. In this case the price of membership is the cost of establishing connections, deterministic private utility is income from work, welfare or crime, social utility reflects the effect of peer-referral on the job market, information or stigma with respect to welfare dependency or crime, and so on, and random private utility has the same interpretation as in the parents-kids example. To further simplify the analysis I will use Brock and Durlauf (2001a) "proportional spillovers" specification, where the label refers to the special multiplicative form of social utility:

$$V_i(g,\omega) = k_g - \rho_g + h_i\omega + J_g\omega m^e_{ig} + \xi_{ig\omega}, \tag{4}$$

[5] This way I'm assuming all social interactions are mediated by others' behavior. This may not be the best way to model relational influences on individual behavior, but has all the practical advantages of an operational definition (see Manski, 2000, for a discussion).

[6] A more interesting parents-kids model actually requires specifying two objective functions, since parents' and kids' preferences are not necessarily convergent: usually parents want to maximize kids' human capital and not the happiness they may derive from "bad" social interactions. However, as will become clear later, the main conclusions of the model I use would hold in a parents-kids model with diverging objectives.

where $J_g > 0$ is a group-specific social interactions parameter. The difference between amenities and price, $k_g - \rho_g$, is the net deterministic private benefit of membership. This specification of social utility captures the following effect: as the majority in the group is expected to chooses $\omega = 1$, so that $m^e_{ig} > 0$, an individual has an incentive to choose $\omega = 1$. In Manski's (1993) terminology, this is an endogenous social effect, as expected mean behavior is determined in equilibrium. Therefore, a specification like (4) is useful to model, for example, social norms that may affect welfare and illegal behavior, or role models and peer effects that may affect investment in human capital[7]. Denoting with p^e_{ijg} individual i's subjective probability that individual j will be a member of group g, with $m^e_{ij|g}$ individual i's subjective expectation of j's behavioral choice given that j is a member of group g, and assuming $I_g > 1$, subjective expected mean behavior that appears in the utility function can be written as the average expected behavior in the group, considering expected group composition:

$$m^e_{ig} = (I_g - 1)^{-1} \sum_{j \neq i} p^e_{ijg} m^e_{ij|g}. \tag{5}$$

As suggested by Brock and Durlauf (2005), the logit structure of their model with exogenous groups can be extended to the two-stage choice process depicted in figure 1 using the nested logit model[8]. In the rest of this section I will describe the parametric assumptions needed to derive such model, so the reader who is familiar with discrete choice models can just browse the new notation I introduce and jump to the next section. There are two equivalent ways of deriving a nested logit model. The more popular one, due to McFadden (1978), is to assume that the vector $(\xi_{iA\omega}, \xi_{iB\omega})$ has a generalized extreme value (GEV) cumulative distribution, such as the following:

[7] However, such specification neglects inter-group interactions, which may be relevant in some applications, as well as funding of local public goods. These aspects are considered, for instance, by Benabou (1993 and 1996), whose models feature cross-group complementarities in production and within-group school funding. The model can be extendend to include local public goods by reinterpreting and endogenizing k_g. For instance, one can write $k_g = f_{Hg}H + f_{Lg}L$, i.e. the salient characteristics of the group are the mean characteristics of the individuals who populate it. This would only reinforce the main theoretical conclusions of the paper.

[8] The nested logit was devised to model situations in which elements of the choice set share observed and unobserved attributes within nests, i.e. groups in this case. See Ben-Akiva and Lerman (1985) and Train (2003) for a detailed exposition.

$$F\left(\xi_{iA\omega},\xi_{iB\omega}\right)=\exp\left(-\sum_{g}\left(\sum_{\omega}\exp\left(-\xi_{ig\omega}\right)\right)^{\beta}\right) \tag{6}$$

where β is the scale parameter. The second, which traces back to Ben-Akiva (1973), is to decompose the model into two interrelated logit models. First, decompose the random component of utility into 3 zero-mean random variables:

$$\xi_{ig\omega}=\varepsilon_{ig\omega}+\varepsilon_{ig}+\varepsilon_{i\omega}, \tag{7}$$

where the last two terms represent the portion of unobserved utility that varies, respectively, across groups and behaviors only. Also assume that (i) the share of payoff due to unobserved elements varying only across behaviors is negligible, i.e. $var\left(\varepsilon_{i\omega}\right)=0$, and (ii) $\varepsilon_{ig\omega}$ and ε_{ig} are independent for all individuals, groups, and behaviors. Such assumptions, of course, are not innocuous, and so deserve a word of comment. The first means that unobserved utility from behavioral alternatives is specific to the social context. The second means that the unobservable propensity to choose a certain behavior given membership is independent of the unobservable preference for the group chosen at the first stage. Whether these assumptions are tenable or not depends on the particular application. For instance, they are plausible in the parents-kids example, since parents may be attached to a group for reasons unrelated to kids' talent, but less so in the work-welfare or crime example, as people less willing to work or more inclined towards crime may tend to join groups which are associated with welfare dependency or high crime rates, in which case the assumption that for a given individual and group $\varepsilon_{ig\omega}$ and ε_{ig} are independent is implausible. Finally, assume that $\varepsilon_{ig\omega}$ is extreme value (EV) distributed, with scale parameter normalized to 1, in which case its cumulative density is

$$G\left(\varepsilon_{ig\omega}\right)=\exp\left(-\exp\left(-\varepsilon_{ig\omega}\right)\right), \tag{8}$$

and that ε_{ig} is distributed such that the maximum of utility with respect to behavior is EV distributed with scale parameter β, i.e. with cumulative density such that

$$\Gamma\left(\max_{\omega}V_i\left(g,\omega\right)\right)=\exp\left(-\exp\left(-\beta\max_{\omega}V_i\left(g,\omega\right)\right)\right). \tag{9}$$

Although apparently more cumbersome, this second way is more useful because it allows one to express directly choice probabilities as the product of marginal and conditional probabilities, which has an immediate interpretation in terms of the two stages of choice (see Train, 2003)[9]. For this reason I will follow the second way and work with the following utility function:

$$V_i(g,\omega) = k_g - \rho_g + h_i\omega + J_g\omega m^e_{ig} + \varepsilon_{ig\omega} + \varepsilon_{ig}. \tag{10}$$

3 Choice

Each individual maximizes utility, i.e. solves:

$$\max_{g,\omega} V_i(g,\omega) \tag{11}$$

subject to[10] $(g,\omega) \in \{A,B\} \times \{-1,1\}$. As illustrated above, this problem is solved in two stages. I will denote with (g_i,ω_i) individual i's choice. At the first stage, each individual chooses a group, given that optimal behavior will be chosen at the second stage, and given beliefs on other individuals' choices over groups, p^e_{ijg}, for each $j \neq i$ and each g:

$$g_i = \arg\max_g \left[k_g - \rho_g + \varepsilon_{ig} + \max_\omega \left(h_i\omega + J_g\omega m^e_{ig} + \varepsilon_{ig\omega}\right)\right]. \tag{12}$$

At the second stage, each individual chooses behavior, given membership in group g, the distribution on others' memberships and expected mean behavior, m^e_{ig}:

$$\omega_i = \arg\max_\omega \left(h_i\omega + J_g\omega m^e_{ig} + \varepsilon_{ig\omega}\right). \tag{13}$$

While individuals choose a specific group and a specific behavior, everybody, ex-ante, can only figure out a probability distribution over others' choices. Such distribution is then used to form expectations about group composition and mean behavior, as required by self-consistent or rational

[9] The two ways are equivalent since the marginals of a GEV distribution are EV.

[10] In order to simplify the analysis, I ignore the budget constraint. The introduction of such constraint would only reinforce the main conclusions I reach, as long as observable types and income are correlated, such as when the individual type is human capital.

expectations. Our parametric assumptions allow the derivation of objective choice probabilities, proceeding backward from the lower level of the decision tree. Since the difference between two EV random variables is logistically distributed, rule (13) implies that the probability individual i chooses behavior ω, given membership in group g, denoted $p_{i\omega|g}$, is equal to the probability that $V_i(g,\omega)$ is greater or equal to $V_i(g,-\omega)$, i.e.

$$p_{i\omega|g} = \frac{\exp\left(h_i\omega + J_g\omega m_{ig}^e\right)}{\exp\left(h_i + J_g m_{ig}^e\right) + \exp\left(-h_i - J_g m_{ig}^e\right)}. \tag{14}$$

In the context of a nested logit model this is known as the "lower model" (see figure 1), and is the model studied by Brock and Durlauf (2001a). So the model I use reduces to theirs when memberships are given. The "upper model", involving rule (12), is simplified by the fact that the maximum of an EV random variable is itself EV distributed. It can be shown (see technical appendix) that the expected value of maximum utility with respect to behavior appearing in equation (12) is simply the log of the denominator in (14), plus Euler's constant (denoted γ):

$$E\max_{\omega}\left(h_i\omega + J_g\omega m_{ig}^e + \varepsilon_{ig\omega}\right) = \log\textstyle\sum_{\omega}\exp\left(h_i\omega + J_g\omega m_{ig}^e\right) + \gamma. \tag{15}$$

Defining $W_{ig}^e \equiv \log\sum_{\omega}\exp\left(h_i\omega + J_g\omega m_{ig}^e\right)$, rule (12) and distribution (9) imply that the probability that individual i chooses group g is

$$p_{ig} = \frac{\exp\left(\beta\left(k_g - \rho_g + W_{ig}^e\right)\right)}{\exp\left(\beta\left(k_A - \rho_A + W_{iA}^e\right)\right) + \exp\left(\beta\left(k_B - \rho_B + W_{iB}^e\right)\right)}. \tag{16}$$

The quantity W_{ig}^e, derived by Ben-Akiva (1973) and known as *inclusive utility*, is a key object since it links upper and lower models. It is the appropriate value of membership, up to amenities, since it is the expected value of the behavioral "choice situation" once in group g. For an individual, the inclusive utility of a group carries to the first stage, where membership is chosen, the relevant information about the endogenous characteristics of that group. These include expected social interactions at the second stage, where behavior is chosen. This point is the key to the whole model, since expected social interactions, through expected behaviors, turn out to be a crucial determinant of group composition, and actual behaviors[11]. The joint probability of being a member of group g and choosing behavior ω is then simply $p_{ig\omega} = p_{ig}p_{i\omega|g}$. Notice that the ratio between such joint choice probabilities for any two alternatives is independent of the attributes of other alternatives within but not across groups. This property is know as *independence from irrelevant nests* (IIN), a relaxation of the more stringent independence from irrelevant alternatives (IIA), which would be an implication of the model if the alternatives in the choice set were not nested[12]. To close the model, we need to specify how the subjective beliefs appearing in equation (5) are formed. Since the model produces objective choice probabilities, rational agents must form their expectations using them, and assume that other agents will do the same. In other words, I am assuming rational expectations, an easy way to close the model: subjective beliefs must be equal to the objective probabilities generated by the model. As for membership, the probability individual i attributes to j being a member of group g must be equal to the probability in (16):

$$p_{ijg}^{e} = p_{jg} \quad \forall i, j. \tag{17}$$

As for behavior, individual beliefs are determined by the mathematical expectation with respect to the binary distribution in (14):

$$m_{ij|g}^{e} = E\left(\omega_j|g\right) = \left(p_{j1|g} - p_{j-1|g}\right) \quad \forall i, j.$$

Therefore, subjective expected mean behavior in group g is equal to the objective mean: $m_{ig}^{e} = m_{ig}$, where

[11]The nested logit model is affected by a complication: choice rules (12)-(13) are consistent with random utility maximization only if certain restrictions on β are imposed. This means we need to restrict β if we want the probability of choosing an alternative to (weakly) increase with the desirable attributes of that alternative. This is equivalent to the requirement that the density of $\varepsilon_{ig\omega} + \varepsilon_{ig}$ be non negative, which is not obvious in general, since the sum of two EV random variables is not itself EV distributed. A sufficient condition for this is $0 \leq \beta \leq 1$, known as the Daly-Zachary-McFadden condition, which has been shown to be unnecessarily strong by Borsch-Supan (1990). A simple necessary and sufficient condition when there are two stages of choice and each nest contains less than four alternatives, has been devised by Herriges and Kling (1996). To keep the model going, I simply assume $0 \leq \beta \leq 1$, which can be tested in empirical applications.

[12]The IIA property generates choice paradoxes such as the red bus/blue bus paradox (Debreu, 1960).

$$m_{ig} = (I_g - 1)^{-1} \sum_{j \neq i} E(\omega_j | g) \, p_{jg}, \tag{18}$$

i.e. the self consistent version of equation (5). When group size is large enough, this expectation is well approximated by the average of individual expected behaviors in the group, m_g:

$$m_g = I_g^{-1} \sum_i E(\omega_i | g) \, p_{ig}. \tag{19}$$

With the imposition of rational expectations all heterogeneity is now due to the individual type. Therefore, we can replace the individual index i with the type index h and rewrite (19) as

$$m_g = \alpha_g^{-1} \sum_{h \in \Theta} f_h E(\omega_h | g) \, p_{hg}. \tag{20}$$

The right hand side of equation (20) will turn out to have a very convenient expression once the memberships market equilibrium condition is imposed. For the moment, we can rewrite choice probabilities in objective form:

$$p_{h\omega|g} = \frac{\exp(h\omega + J_g \omega m_g)}{\sum\limits_{\omega=-1,1} \exp(h\omega + J_g \omega m_g)} \tag{21}$$

$$p_{hg} = \frac{\exp\left(\beta\left(k_g - \rho_g + W_{hg}\right)\right)}{\sum\limits_{\nu=A,B} \exp\left(\beta\left(k_\nu - \rho_\nu + W_{h\nu}\right)\right)}, \tag{22}$$

where

$$W_{hg} = \log \sum_{\omega} \exp(h\omega + J_g \omega m_g). \tag{23}$$

4 Equilibrium

In this economy an equilibrium is defined in the standard way: a set of prices $\{\rho_A^*, \rho_B^*\}$, a set of expected mean behaviors $\{m_A^*, m_B^*\}$, and a set of membership probabilities $\{p_{hA}^*, p_{hB}^*\}_{h \in \Theta}$, such that the memberships market clears and, for each individual, expectations are self-consistent and utility is maximized. In this section I analyze in detail these three components of the equilibrium and their interplay. I refer to them, respectively, as *memberships market equilibrium*, *behavioral equilibrium*, and *sorting equilibrium*[13].

4.1 Memberships market equilibrium

The expected demand for membership in any group g is equal to the sum of the individual choice probabilities over that group, given by (22). On the other side of the market, the supply is fixed, and equal to I_g. Walras' law implies we can focus on one of the two groups only, say group A whose relative capacity is α_A. Therefore, market clearing requires:

$$f_H p_{HA} + f_L p_{LA} = \alpha_A \tag{24}$$

Proposition 1 *A unique set of membership prices, $\{\rho_A^*, \rho_B^*\}$, solving equation (24) exists, for any pair of expected mean behaviors, $\{m_A, m_B\}$.*

Proof. See Technical Appendix.

As one might expect, the market clearing difference between membership prices depends on social interactions, i.e. contains a *social premium.* This is easily derived in the special case[14] of types equally distributed in the population ($f_H = f_L = \frac{1}{2}$) and groups of equal capacity ($\alpha_A = \alpha_B = \frac{1}{2}$). In this case, replacing (22) into (24) and solving for the market clearing difference in prices, we have

$$\rho_A^* - \rho_B^* = (k_A - k_B) + \tfrac{1}{2}\left[(W_{HA} - W_{HB}) + (W_{LA} - W_{LB})\right]. \tag{25}$$

The first term on the right hand side reflects the difference in amenities, and the second term is the social premium. The latter reflects the different value of social interactions in the two groups for the two types[15]. In the

[13] Of course this separation is fictitious and put forth for the sake of exposition, as the equilibrium is determined as a whole.

[14] The general case involves a quadratic whose roots cannot be recovered analytically.

[15] When individuals do not interact, i.e. $J_A = J_B = 0$, or social interactions have the same utility in the two groups, i.e. $J_A m_A = J_B m_B$, inclusive utilities in (25) cancel out and the social premium is zero.

special case that leads to (25), the social premium is the average difference between inclusive utilities of groups A and B for the two types, or the average evaluation of how better individuals will do in group A compared with B. This in turn reflects the average evaluation of social interactions in the population. It may appear odd that this average uses equal weights, since different types have different propensities to choose different groups. The reason is that when the proportion of types and group size is the same, the two groups are ex-ante (before sorting takes place and behavior is chosen) identical, up to different amenities. As can be seen rearranging the terms in the social premium above, a group commands a positive social premium if it offers, on average across types, the "best interactions", which are associated with the highest value of the choice situation in the two groups, i.e. $\max\{(W_{HA} + W_{LA}), (W_{HB} + W_{LB})\}$. Referring to the parents-kids example, a school will charge higher tuitions if it has good teachers (high k_g) but also if it offers a social environment that favors effort, for instance through peer-pressure or peer-aid, to students from any kind of family. As I will show in the next section, the social premium plays a critical role in sustaining stratified groups. This is a standard feature of sorting models with social effects. The boundary values of the price difference with respect to the proportion of H-types in the population, derived for general group size using equation (24), offer useful insights into the social premium:

$$\left(\rho_A^* - \rho_B^*\right)|_{f_H=0} = \beta^{-1}\log\frac{\alpha_B}{\alpha_A} + (k_A - k_B) + (W_{LA} - W_{LB}) \quad (26)$$

$$\left(\rho_A^* - \rho_B^*\right)|_{f_H=1} = \beta^{-1}\log\frac{\alpha_B}{\alpha_A} + (k_A - k_B) + (W_{HA} - W_{HB}) \quad (27)$$

The first term on the right hand side is a *size premium*, normalized by the scale of utility, and depends on relative group size only. Such premium exists because of the stochastic utility of membership, since there is a positive probability that many individuals have a high evaluation for the smaller group. The last term is again the social premium. When the population is homogeneous, it reduces to the difference between inclusive utilities across groups, for the only type. It is tempting to think of the social premium as the weighted average across types of differences in inclusive utilities, but in the general case the relation between equilibrium price differential and proportions of types in the population is not linear. However, a simple thought experiment can establish such relation is weakly monotonic. Consider a sorting equilibrium with $f_H = 0$. Now introduce an H-type individual and withdraw an L-type one. The group the H-type chooses with higher probability becomes more attractive to everybody and so the social premium rises in absolute value. If the H-type instead is equally likely to choose the two groups, the social premium does not move. At the new equilibrium, repeat the thought experiment until $f_H = 1$. This proves that the social premium moves monotonically between the extremes $W_{LA} - W_{LB}$ and $W_{HA} - W_{HB}$. When $\beta = \infty$, the price differentials (26)-(27) reduce to the equilibrium differentials in Benabou's (1993) model, where utility is deterministic and equilibria are fully stratified. Therefore, since the variance of an EV distribution goes to zero as the scale parameter increases, this model reduces to his when we eliminate randomness in group preference, i.e. shut off ε_{ig}.

4.2 Behavioral equilibrium

The number of h-type individuals demanding membership in group g is $I f_h p_{hg}$. Dividing this by group size, we obtain the proportion of h types demanding membership in the group, $\alpha_g^{-1} f_h p_{hg}$. When equilibrium condition (24) holds, such fraction determines the distribution of types in group g:

$$f_{hg} = \alpha_g^{-1} f_h p_{hg} \tag{28}$$

Substituting this expression into (20), expected mean behavior in group g can be conveniently expressed as[16]

$$\begin{aligned} m_g &= \sum_{h \in \Theta} f_{hg} E\left(\omega_h | g\right) \\ &= \sum_{h \in \Theta} f_{hg} \tanh\left(h + J_g m_g\right). \end{aligned} \tag{29}$$

A behavioral equilibrium in group g is a fixed point of such equation, i.e. a value m_g^* solving:

[16] The second equality derives from the convenient expression of the conditional expectation in the binary model, $E\left(\omega | g\right) = p_{h1|g} - p_{h-1|g}$, and uses the definition $\tanh\left(X\right) = \frac{e^X - e^{-X}}{e^X + e^{-X}}$.

$$m_g^* = f_{Hg} \tanh\left(H + J_g m_g^*\right) + (1 - f_{Hg}) \tanh\left(L + J_g m_g^*\right) \quad (30)$$

Since the tanh function is continuous and $m_g : [-1, 1] \rightarrow [-1, 1]$, existence of a behavioral equilibrium is a direct consequence of Brouwer's fixed point theorem:

Proposition 2 *For any arrangement of individuals across the two groups, there exists an equilibrium set of expected mean behaviors, $\{m_A^*, m_B^*\}$.*

This is the extension of one of the results in Brock and Durlauf (2001a) to the case of two types and two endogenous groups. Since H and L are positive, the following two propositions, illustrated in figure 2, exhaust the possible behavioral equilibrium configurations[17].

(U) *Equation* (30) *has a unique root, $m_{g,1}^* > 0$.*

(M) *Equation* (30) *has three roots, $m_{g,1}^* > 0$, $m_{g,2}^* < 0$* and *$m_{g,3}^* < 0$.*

Two additional results which are central to the analysis extend to this model: in case of multiple equilibria the intermediate one is dynamically unstable[18], and the stable equilibria are Pareto-rankable[19]. Here, since H and L are positive, the best equilibrium is the positive root of (30).

[17]Equation (30) cannot have more than three roots. Notice that the second derivative of the function on the right hand side of (30) changes sign at most once over the interval $[-1, 1]$, i.e. the function has at most one inflection point, which implies equation (30) has at most three roots.

[18]Imagine that at equilibrium one perturbs behavioral choices. The adjustment process can be thought of as governed by the dynamic version of equation (30):

$$m_{g,t}^* = f_{Hg} \tanh\left(H + J_g m_{g,t-1}^*\right) + (1 - f_{Hg}) \tanh\left(L + J_g m_{g,t-1}^*\right)$$

and the same argument in Brock and Durlauf (2001a) applies.

[19]This is easy to verify using inclusive utility (equation 23), which expresses utility before behavior is chosen and so is the appropriate welfare index (I will return on this point in section 6). When $H = L = 0$, inclusive utility is symmetric about zero, and so the two extreme equilibria are Pareto-equivalent. But since H and L are positive, its value at $m_{g,1}$ is larger than at $m_{g,3}$ for both types, and $m_{g,1}$ is Pareto-superior.

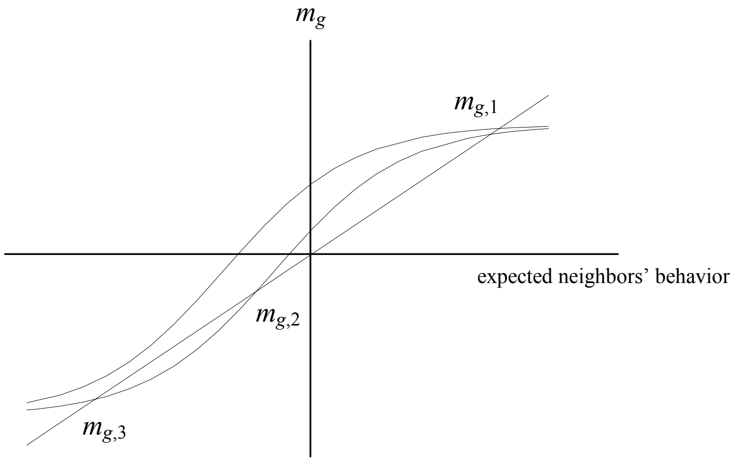

Figure 2. Unique and multiple behavioral equilibria.

What this work elucidates in addition is that the possibility of multiple equilibria not only depends on the relation between private and social utility, but also on the sorting process when the population is heterogeneous. This link between multiple equilibria and choice of membership is immediate from equation (30), since the fraction of H-types in the group, f_{Hg}, is endogenously determined. This can be stated formally as follows.

Proposition 3 *Suppose J_g, H and L are such that $f_{Hg} = 1$ implies (U) and $f_{Hg} = 0$ implies (M). Then there exists a unique threshold $\widetilde{f}_H$ such that if $f_{Hg} > \widetilde{f}_H$, then (U) holds, and if $f_{Hg} < \widetilde{f}_H$, then (M) holds.*

Proof. See Technical Appendix.

In words, proposition 3 states the following. Suppose that the model parameters are such that a group populated only by H types has a unique behavioral equilibrium, while a group populated only by L types has multiple behavioral equilibria. That is, when H types are segregated, their private incentives to choose $\omega = 1$ are stronger than social pressure, for any possible behavior chosen by their neighbors, while when L types are segregated there is a range of neighbors' mean behavior over which the reverse is true. Then, there exists a unique critical mixed group composition, generating a unique stable behavioral equilibrium, such that if we marginally increase the fraction of H types then multiple stable equilibria appear. Referring again to the parents-kids example, this means that in the presence of endogenous neighborhood effects it is possible that when enough H-type households (whose kids have stronger private incentive not to engage in risky behavior) move away from a neighborhood, this may end up in a bad behavioral equilibrium - for instance high rates of school dropout, teenage pregnancy, or youth crime - even if the neighborhood was previously in a good state, as kids' social incentives to behave in a socially desirable way drop while their private incentives are given. The characterization of the sorting equilibrium will allow us to establish under which conditions this kind of transition may endogenously occur.

4.3 Sorting equilibrium

The third component of equilibrium, sorting equilibrium, is trickier that the previous two, because of the possibility of multiple behavioral equilibria. This possibility implies that inclusive utility of membership in a group is not uniquely determined. This can be seen in equation (23): if more than one fixed point exists, then W_{hg} can assume more than one value in equilibrium. Since the model does not suggest which behavioral equilibrium will be selected, the natural thing to do is to consider beliefs over behavioral equilibria in the two groups before sorting takes place, and then use expected inclusive utility to find a sorting equilibrium. It is possible to generate such beliefs reinterpreting the model in a dynamic sense. Imagine we allow individuals, at random times, to revise their behavior in response to realized mean behavior in the group. Then, as shown in Blume and Durlauf (2003), when the population is large enough the invariant distribution of the stochastic process generated by such revisions converges to a distribution on the stable equilibria generated by equation (30). In other words, the stable equilibria of the static model become basins of attraction in a dynamic specification. The beliefs of rational individuals over behavioral equilibria should agree with such distribution, i.e. the dynamics introduce a form of coordination[20]. Using this

[20] Of course when multiple equilibria exist but this dynamic reinterpretation is not appropriate, for instance because in a particular application choices cannot be revised, individuals beliefs can be anything and it's impossible to pin down both behavioral and sorting equilibria, unless one introduces some other coordination mechanism. In many applications of social interactions models, coordination mechanisms may be easy to find: after all, the hallmark of sociality is that people are able to coordinate. For instance, if a school is known for the education achievements of its former students, the high equilibrium

argument, I proceed under the assumption that beliefs over behavioral equilibria are represented by a common distribution $\mu\left(m^*_{g,z}\right)$, where $z = 1, 2, 3$ indexes the three possible equilibria represented in figure 2. Notice that since the intermediate equilibrium is dynamically unstable, the previous argument implies $\mu\left(m^*_{g,2}\right) = 0$: individuals believe the system will not remain for a long time at an unstable equilibrium. Under these assumptions, equilibrium inclusive utility can be rewritten in expected utility form:

$$W^*_{hg} = \sum_z \mu\left(m^*_{g,z}\right) \log\left[\exp\left(h + J_g m^*_{g,z}\right) + \exp\left(-h - J_g m^*_{g,z}\right)\right] \tag{31}$$

Since each behavioral equilibrium solves equation (30), each $m^*_{g,z}$, using equation (28), is an implicit function of the endogenous variable p_{hg}, for any h. Using this fact into (31), we can rewrite equation (22) as:

$$p_{hg} = \frac{\exp\left(\beta\left(k_g - \rho^*_g + W^*_{hg}\left(p_{hg}\right)\right)\right)}{\sum\limits_{\nu=A,B} \exp\left(\beta\left(k_\nu - \rho^*_\nu + W^*_{h\nu}\left(p_{h\nu}\right)\right)\right)} \tag{32}$$

This is the equilibrium probability function over individual membership. Since $p_{hA} + p_{hB} = 1$, the right hand side of equation (32) can be seen as a function of p_{hg} only, mapping the interval $[0, 1]$ into itself. Therefore, existence of a sorting equilibrium follows directly from Brouwer's fixed point theorem:

Proposition 4 *Given a pair of probability distributions $\mu\left(m^*_{g,z}\right)$, $g = A, B$, over the set of behavioral equilibria, a set of equilibrium choice probabilities over groups, $\{p^*_{hA}, p^*_{hB}\}$, exist for each type h.*

Given the distribution of types in the population, a sorting equilibrium determines an equilibrium distribution of types across groups, $\{f^*_{hA}, f^*_{hB}\}$, according to equation (28). Therefore, based on propositions 1, 2 and 4, we can establish the existence of a general equilibrium:

Proposition 5 *An equilibrium, i.e. a triplet $\{\rho^*_A, \rho^*_B\}$, $\{p^*_{hA}, p^*_{hB}\}_{h\in\Theta}$, and $\{m^*_A, m^*_B\}\}$, exists in the binary choice model with two endogenous groups.*

is a focal point, in which case it should be assigned probability 1.

In the next two sections I explore three questions which are central in any sorting model, and that elucidate further the feedback between endogeneity of memberships and social influences. The three question, in the order I address them, are: (i) is there a tendency towards stratification along the individual type? (ii) is the sorting equilibrium unique? (iii) Is the decentralized equilibrium efficient? Section 5 addresses the first two, showing that the answers are respectively "yes" and "no", and section 6 addresses the third, showing that the answer depends on the parameters.

5 Characterizing sorting equilibria

I will tackle the stratification question first, defining the economy integrated, or not stratified, when the composition of groups and population are the same, i.e. $f_{hg} = f_h$, which in turn implies $p_{hg} = \alpha_g$, for all g and h. At the other extreme, the economy is said to be segregated, or fully stratified, when $f_{hg} = 0$ for at least one type h and one group g, which implies $p_{hg} = 0$. Any intermediate case is associated with some degree of stratification. Since there are only two groups and two types, we can define a stratification index for this economy focusing on a single type, as $\max_g |f_{Hg} - f_H|$, which lies in the interval $[0, f_H]$. Since membership probabilities map into such an interval, a convenient way to proceed is to define, for each type h, *isoprobability curves* in the m_g-ρ_g space. This is a natural extension of bid-rent analysis, as used for instance in Benabou (1996), to a random utility setting. Isoprobability curves are sets of price-expected mean behavior combinations such that membership probability is constant, with higher loci associated with lower probability levels. The slope of a curve at any point is the marginal rate of substitution between membership price and mean behavior at that point, for a certain probability level. This has a simple expression in the binary case:

$$MRS^h_{m\rho} = -\frac{\partial p_{hg}/\partial m_g}{\partial p_{hg}/\partial \rho_g} = 2J_g \tanh\left(h + J_g m\right) \tag{33}$$

In the presence of positive social interactions ($J_g > 0$) such a ratio is increasing in h, i.e. isoprobability curves exhibit the single crossing property, as illustrated in figure 3.

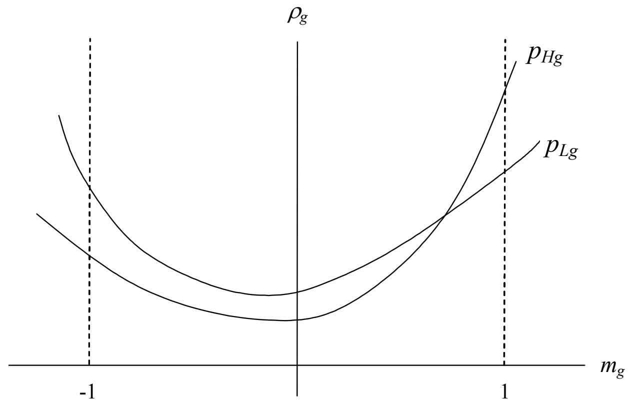

Figure 3. Isoprobability curves of the two types.

In the figure, two curves for types H and L and one group are sketched. When the behavioral equilibrium is unique, the intersection at equilibrium probability levels locates a general equilibrium, since it's the only point where individuals in the same group face identical price and expected mean behavior. However, such graphical interpretation of the equilibrium is not possible when multiple behavioral equilibria exist[21]. In either case isoprobability curves represent preferences: when they are U-shaped in the interval $[-1, 1]$, which happens when H and L are small enough, a ratio increasing in h means that the H-types have a higher willingness to pay than the L-types to be members of a group with higher expected mean behavior at high levels of m_g. However, at low levels, agents are actually willing to pay for *lower* expected mean behavior. This happens because social effects and private incentives work in opposite directions and the former offsets the latter, in which case the L-types are willing to pay more, since their private incentives are the weakest. On the other hand, when H and L are large enough, isoprobability curves are increasing over $[-1, 1]$ and the H-types always have a higher willingness to pay for membership in high expected mean behavior groups. Single crossing is sufficient for equilibrium stratification, since types are publicly observed. Expression (33) also shows that the slope of isoprobability curves increases in the magnitude of social interactions, J_g: the more important interactions are in a group, the more individuals are willing to

[21] The reason is that in this case individuals face more than one level of expected mean behavior in equilibrium, while price is still uniquely determined by equation (24), using expected inclusive utility.

pay for extreme values of expected mean behavior in that group. This in turn will affect the degree of stratification in equilibrium. In terms of the parents-kids example, in this model better educated parents have a higher willingness to pay for memberships in neighborhoods and schools with higher expected mean behavior, which offer interactions for their kids that favor effort and so high human capital. Since education and income are positively correlated, this is sufficient to generate rich and poor neighborhoods, and schools for well off and less well off families. Such separation will become more pronouced the more interactions are important for performance. I will now turn to question (*ii*): how many sorting equilibria exist?

Proposition 6 *In presence of positive social interactions ($J_g > 0$), there exist three sorting equilibria, for any initial arrangement of the population into groups. One of them is integrated, and is unstable. The other two are stratified by some degree, and are stable. Furthermore, the degree of equilibrium stratification increases with the interactions parameter, J_g, and decreases with the dispersion of preferences over memberships, β^{-1}.*

This result is well known in the endogenous neighborhoods literature that focuses on two groups, e.g. Benabou (1993 and 1996) and Becker and Murphy (2000). The fact that social interactions are a stratifying force when individuals can choose their reference group, and complete stratification along the individual type are recurring results in these models. However, segregation is not a necessary condition for equilibrium here: because of the stochastic part of utility, in equilibrium groups are stratified but imperfectly so. Heterogeneity along more than one dimension (observable type and unobservable preferences here) is sufficient for imperfect stratification, as shown, for instance, in Epple and Platt (1998). In addition to being more realistic, this feature of the model offers two advantages. First, it is possible to do comparative statics on the degree of equilibrium stratification. For instance, proposition 6 says that the equilibrium degree of stratification is small if either of the following holds: (i) $J_g \to 0$ (i.e. interactions are weak): when the endogenous social effect is negligible, group composition does not affect utility much, and so the incentive to segregate is weak; (ii) $\beta \to 0$ (i.e. maximum utility with respect to behavior has high dispersion): individuals have a possibly strong unobservable preference for a certain group, and this component, rather than prices and expected mean behavior, drives membership probabilities. The second advantage is that the model is directly useful for empirical purposes. The rest of the section is devoted to proving proposition 6: this offers further insights into the properties of the model.

Consider the integrated economy first. As illustrated above, this is defined by $p_{hg} = \alpha_g$, for both types and groups, which implies $m_A = m_B$. In order for this to be an equilibrium we need to find a pair of prices consistent with the following two properties: different types have equal probabilities of being members of a given group, and expected mean behavior is equal across groups. Notice that the latter implies $W_{hA} = W_{hB}$. Therefore, no group can command a social premium, and it is enough to choose a pair of prices such that their difference only reflects different amenities and size. So integration is an equilibrium.

However, it is unstable if we introduce some dynamics. To see this, it is sufficient to perform the following thought experiment. Start at integration, and perturb the equilibrium so that, for some small ϵ, $p_{HA} = \alpha_A + \epsilon$ and $p_{HB} = \alpha_B - \epsilon$, i.e. the H-type now has a higher probability of being member of group A and a corresponding smaller probability of being member of group B. Now allow individuals to revise their membership choice. All we have to do is to show that the system will not revert to the integrated equilibrium. Following the perturbation, the fraction of H-types in group A and of L-types in B, f_{HA} and f_{LB}, increase according to equation (28). Since $H > L$, the properties of the *tanh* function imply that the value of the stable roots of (30) move in the same direction f_{HA} moves[22], so the equilibrium value (or values) of m_A and m_B, respectively, increases and decreases. This implies that expected inclusive utility increases in group A and decreases in group B for both types[23]. This generates a positive social premium on the memberships market, but because of single crossing the implied p_{HA} and p_{LA} diverge. So membership probabilities cannot revert to the integrated equilibrium where p_{HA} and p_{LA} are equal. The adjustment process will continue, repeating these steps, with f_{HA}, m_A and p_{HA} increasing. Since expected mean behavior and membership probabilities are bounded, the system must reach a new equilibrium with stratification. The adjustment process in group B is symmetric, with the notable difference that since f_{HB} is falling, mul-

[22] The reader can convince herself or himself considering the proof of proposition 3 and drawing a few tanh fucntions along with the 45 degree line.

[23] This is obvious if the behavioral equilibrium in either group is unique, and follows from the fact that H and L are positive if multiple equilibria exist (since in case of multiplicity inclusive utility varies more when evaluated at the positive m_g than at the negative m_g) and from the assumption $\mu\left(m^*_{g,2}\right) = 0$.

tiple behavioral equilibria may appear in group B even if the equilibrium was unique under integration, which causes a discontinuous drop in inclusive utility of group B and a rapid increase in stratification. The properties of the logistic function in equation (32) imply that we cannot have more than three equilibria, so if we repeat the experiment starting at full stratification ($f_{HA} = 1$) the same equilibrium is necessarily reached, which establishes stability. The existence of another stable equilibrium with stratification follows immediately if we invert the roles of groups A and B in the thought experiment[24]. This completes the proof of the first part of proposition 6, whose meaning is illustrated in figure 4, where the two sides of equation (32) are depicted. So, the two stable equilibria can be thought of as the steady states of a dynamic adjustment process when we start from arbitrary group compositions. To return to the previous example, as an initial small group of rich families move out of an integrated neighborhood, a cascade is triggered with more and more rich families leaving in search of better interactions for their kids, until neighborhoods are stratified.

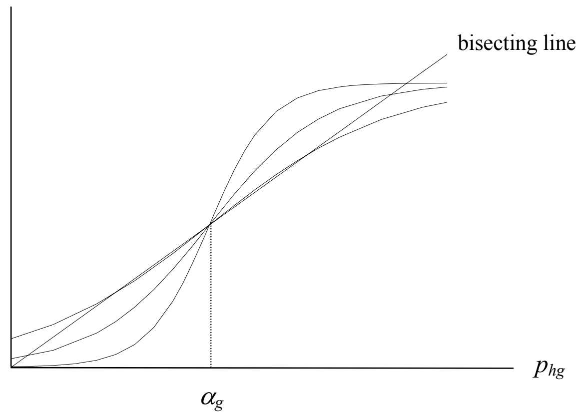

Figure 4. Fixed points of the membership probability function

The different curves in figure 4 correspond to different preference parameters. Specifically, steeper curves, which generate higher degrees of equilibrium stratification, are associated with a larger value of the interactions parameter, J_g, and with smaller dispersion of preferences over memberships, i.e. larger

[24]The two stratified equilibria are not symmetric in general, since amenities may differ, i.e. $k_A \neq k_B$.

β, as can be seen on the right hand side of (32), which completes the proof of proposition 6. In particular, stratification is maximal as $\beta \to \infty$, and so the sorting equilibrium in Benabou (1993 and 1996) and Durlauf (1996) is the limiting case of this model.

6 Welfare analysis

This section analyzes question (*iii*): is the decentralized equilibrium efficient? To answer, we must figure out how a social planner with an objective function reflecting individual preferences would choose group compositions and subsequent behavior. The planner internalizes both first stage externalities (when choosing a group, individuals do not take into account the effect of their choice on others' inclusive utility) and second stage externalities (when choosing behavior, individuals do not take into account the effect of their choice on others' social utility). Like the decentralized problem, the planner's problem can be decomposed into two stages: first choose behavior as a function of group composition, then choose group composition. The first of these choices was analyzed by Brock and Durlauf (2001a) under the assumption that planner's objective function is itself extreme value distributed. Under the specification of social utility I use here, i.e. proportional spillovers, the planner chooses the positive root of the single-type analog of equation (30) when the interactions coefficient is $2J_g$. I will denote m^o this root, and in general use superscript o for variables evaluated at the social optimum. Therefore, subsidizing the choice of $\omega = 1$ at the second stage, so that social utility is doubled, allows to decentralize the planner's solution.

Given that we can decentralize efficient behavior, how should the planner choose group composition? The appropriate measure of social welfare with respect to group composition is the sum of equilibrium inclusive utilities and amenities, i.e. the aggregate value of memberships before behavior is chosen and given that each individual will behave optimally under the subsidy at second stage. This criterion is the stochastic equivalent of Benabou's (1993 and 1996), who considers aggregate surplus generated by the two communities, since inclusive utility can also be interpreted as expected individual surplus from membership. So the planner maximizes $\sum_h \sum_g k_g + W_{hg}\left(m_g^o\right)$ with respect to group composition, or equivalently, dropping amenities (which are constant across types) and simplifying notation, the planner solves:

$$\max_{f_{HA}} f_{HA} W^o_{HA} + (1 - f_{HA}) W^o_{LA} + f_{HB} W^o_{HB} + (1 - f_{HB}) W^o_{LB}$$

$$\text{s.t.} \qquad \alpha_A f_{HA} + \alpha_B f_{HB} = f_H.$$

The first order necessary condition for a maximum reduces to:

$$\begin{aligned} & I_A^{-1}\left[W^o_{HA} - W^o_{LA} + \left(f_{HA}\frac{\partial W^o_{HA}}{\partial m_A} + (1 - f_{HA})\frac{\partial W^o_{LA}}{\partial m_A}\right)\frac{\partial m^o_A}{\partial f_{HA}}\right] \\ = \; & I_B^{-1}\left[W^o_{HB} - W^o_{LB} + \left(f_{HB}\frac{\partial W^o_{HB}}{\partial m_B} + (1 - f_{HB})\frac{\partial W^o_{LB}}{\partial m_B}\right)\frac{\partial m^o_B}{\partial f_{HB}}\right] \end{aligned} \tag{34}$$

This condition balances across groups the effect on total inclusive utility of a 1% increase in the fraction of H-types, normalized by group size. What does this imply for optimal group composition? First of all, notice that since the optimal m is unique and has the same sign as type, the left hand side is increasing and the right hand side is decreasing in f_{HA}, and so optimal group composition is unique. Next, consider a simple intelligible case: suppose groups have equal capacity, $I_A = I_B$, and equal intensity of interactions, $J_A = J_B$. Under these conditions, one immediately sees that the integrated economy satisfies (34)[25]. So in this case, integration is the optimal arrangement and the decentralized equilibrium is inefficient. Now suppose $I_A > I_B$. Then efficiency requires a larger fraction of H-types in group A, i.e. the planner moves individuals with the strongest private incentives to the largest group, where they can benefit more people via social interactions. If on the other hand $J_A > J_B$, since the marginal effect of expected mean behavior increases in the interactions parameter, the planner optimally moves some H-types to the group with weakest interactions. The intuition is that fewer H-types are needed to provide positive social effects if interactions are strong enough. In the general case, we can conclude that decentralized decision-making generates inefficiently large stratification whenever groups are similar enough. When groups differ considerably, al-

[25] Under integration, inclusive utilities are equal for a given type across groups, and under equal interactions they are equal at the margin as well.

though the decentralized equilibrium is generally inefficient, we cannot say whether stratification is too high or too low. A special and interesting case in which the equilibrium is inefficient occurs when the planner has powers limited to group composition. An example is a school board that can choose class composition but cannot subsidize students' effort in the presence of peer effects. In this case multiple behavioral equilibria are possible again, and decentralized sorting will lead to inefficiently high stratification whenever the parameters are such that the behavioral equilibrium is unique under integration but multiple equilibria appear under equilibrium stratification. The reason is that expected inclusive utility drops discontinuously in a group as multiple equilibria appear.

7 Econometrics

The reason why endogenous memberships are a concern in empirical social interactions work is that - in absence of genuinely randomized experiments - individuals plausibly self-select on unobservables into the groups that compose the sample. This violates the orthogonality condition between error and regressors required in regression analysis. The natural solution to this problem is to use instrumental variables. However, as discussed in Blume and Durlauf (2006), valid instruments are usually hard to find, and it is preferable to model self-selection explicitly within the statistical analysis. In this section I will suggest two ways the theoretical framework developed in this paper can be useful to model self-selection, based on general equilibrium[26]. First, notice that the choice model of section 3 is tantamount to a sample selection model. Referring to the notation introduced in that section, the behavioral equations, for given group g_i chosen optimally at first stage, are:

$$\begin{aligned} y^*_{i\omega} &= V_i(g_i, 1) - V_i(g_i, -1) \\ y_{i\omega} &= 1\left[y^*_{i\omega} > 0\right], \end{aligned} \tag{35}$$

where $1[A]$ is the indicator function assuming value one if A is true and zero otherwise, and $\omega = 1$ if and only if $y_{i\omega} = 1$. The selection equations are:

[26] There are several alternatives to regression analysis to test the presence of social interactions, and good reasons to pursue them. However, this does not mean regression models are useless, especially if one considers the availability of increasingly detailed microdata.

$$
\begin{aligned}
y_{ig}^{*} &= \max_{\omega} V_i(g,\omega) - \max_{\nu \neq g} \max_{\omega} V_i(\nu,\omega) \\
y_{ig} &= 1\left[y_{ig}^{*} > 0\right],
\end{aligned} \tag{36}
$$

with individual i observed in group g if and only if $y_{ig} = 1$. Dubin and Rivers, (1989/1990) have studied this model as a way to eliminate selection bias in discrete choice settings. It turns out that the log likelihood function they derive from a model like (35)-(36) is equivalent to the log likelihood function generated by the nested logit model, i.e.

$$
L = \sum_i \sum_\omega y_{i\omega} \log p_{i\omega|g_i} + \sum_i \sum_g y_{ig} \log p_{ig}, \tag{37}
$$

which suggests a fully parametric empirical route. We first need to generalize the theoretical framework developed above. For the sake of tractability, this was based on binary choice, two groups and two types. In order to construct a useful econometric model, I will stick to binary choice for behavior ($\omega = -1, 1$) but allow multinomial choice for group ($g = 1, ..., G$) and a continuum of types, summarized in a vector of individual characteristics X_i. Also consider a vector of group-specific variables, Y_g, previously collapsed into the k_g index, and allow these to affect behavior as well as membership. Finally, sample size is normally such that we need to assume a unique J common to all groups. Then, introducing a constant a and vectors of coefficients c, d, b and δ, equations (21)-(23) can be rewritten as follows:

$$
p_{i\omega|g} = \frac{\exp\left(a\omega + c'\omega X_i + d'\omega Y_g + J\omega m_g\right)}{\sum_w \exp\left(aw + c'wX_i + d'wY_g + Jwm_g\right)} \tag{38}
$$

$$
p_{ig} = \frac{\exp\left(\beta b \rho_g + \beta\delta' Y_g + \beta W_{ig}\right)}{\sum_\nu \exp\left(\beta b \rho_\nu + \beta\delta' Y_\nu + \beta W_{i\nu}\right)} \tag{39}
$$

$$
W_{ig} = \log \sum_\omega \exp\left(a\omega + c'\omega X_i + d'\omega Y_g + J\omega m_g\right) \tag{40}
$$

The problem is to maximize (37) with respect to $(a, b, c, d, \delta, J, \beta)$ subject to (38)-(40). The fully parametric character of this selection correction procedure may be unsatisfactory. However, since it maps into an equilibrium model, it can be useful to perform a robustness test in applied work involving binary choice and possibly self-selection[27].

Second, it is possible to exploit the theory to correct semiparametric linear models, which are popular tools in the empirical social interactions literature. For instance an increasing number of applied studies uses the linear probability model to identify social interactions. Such model approximates linearly the response of an underlying binary choice $\omega = 0, 1$. In terms of the notation used here the model is:

$$\Pr_i(\omega = 1) = a + c'X_i + d'Y_g + Jm^e_{ig} + \varepsilon_i. \tag{41}$$

As well known, this model suffers from several problems[28]. The most popularized is what Manski (1993) labeled the reflection problem, a special case of non identification due to the possible collinearity between contextual effects and mean individual effects[29]. A second difficulty is the selection problem: in presence of sorting the orthogonality condition $E\left(\varepsilon_i | X_i, Y_g, m^e_{ig}\right) = 0$ required for (41) to be a valid regression equation is violated. A third major problem is the possibility of unobserved group effects, which can be solved using panel data, but I which I will not consider here[30]. Brock and Durlauf (2001b) show that reflection and selection problems can be solved at once modelling selection into groups. Essentially, if individual i has chosen group g according to some selection rule of the kind $R_i\left(X_i, Y_g, m^e_{ig}, \varepsilon_i\right) \geqslant 0$, then model (41) must be corrected using an estimate of $E\left(\varepsilon_i | R_i \geqslant 0\right)$ as an additional regressor, a procedure pioneered by Heckman (1979). The theoretical model developed in this paper provides such selection rule in equation (12), and so allows to construct a correction term based on equilibrium behavior. Denoting with g_i the group optimally chosen in equilibrium by individual i, the selection rule was:

$$g_i = \arg\max_g \max_\omega V_i(g, \omega) = \arg\max_g \left(k_g - \rho_g + W_{ig} + \varepsilon_{ig}\right). \tag{42}$$

[27] The nested logit model can be easily estimated by maximum likelihood using common statistical packages.

[28] See Brock and Durlauf (2001b) and Blume and Durlauf (2005) for an exhaustive exposition.

[29] The problem can be seen imposing the self-consistency condition $m^e_{ig} = E(\omega)$ and rewriting (41) in reduced form.

[30] See Blume and Durlauf (2005).

So, using the notation I introduced above, the correction term is:

$$E\left(\varepsilon_i|\beta b\rho_g + \beta\delta' Y_g + \beta W_{ig} \geqslant \eta_g\right), \tag{43}$$

where

$$\eta_g = \max_{\nu\neq g}\left(\beta b\rho_\nu + \beta\delta' Y_\nu + \beta W_{i\nu} + \varepsilon_{i\nu}\right) - \varepsilon_{ig}. \tag{44}$$

Following Lee's (1983) extension of Heckman's (1979) procedure to the multinomial case, this term can be estimated parametrically in two steps. First, estimate model (38)-(40) and use estimates $\widehat{\beta b}$, $\widehat{\beta\delta}$, and $\widehat{\beta W_{ig}}$ to estimate p_{ig} in (39). Second, assume the error ε_i in equation (41) is normally distributed with standard error σ, and denote with r_g the correlation between $\sigma\varepsilon_i$ and the transformed random variable $\Phi^{-1}\left(\max_{\nu\neq g}\max_\omega V_i\left(g,\omega\right) - \varepsilon_{ig_\nu}\right)$, where Φ is the standard normal cdf. The estimate of the conditional expectation in (43) is $r_g\sigma\phi\left(\Phi^{-1}\left(\widehat{p_{ig}}\right)\right)\left(\widehat{p_{ig}}\right)^{-1}$, where ϕ is the standard normal pdf. As shown in Brock and Durlauf (2001b), as long as r_g is not zero, using this term in equation (41) eliminates selection-bias and allows identification. This procedure integrates the two approaches to modelling selection proposed by Brock and Durlauf (2005), but of course is fragile because of its reliance on parametric assumptions. If an estimate of inclusive utility of groups can be obtained (for instance using the method of Ioannides and Zabel, 2004), then various semiparametric methods can be applied to obtain an estimate of (43). Newey et al. (1990) discuss and use two such methods: one relies on estimating the distribution of η_g using kernel methods, and the other on estimating a series approximation of the correction term (43).

8 Conclusion

In this paper I have analyzed a model of choice subject to endogenous social interactions, when memberships into the groups exerting the social effect are also endogenous, and individuals differ in observables as well as unobservables. The main findings concern: (1) the equilibrium relation between expected social interactions and expected group composition, which implies imperfect and inefficient stratification sustained by a social premium on the memberships market, and (2) the econometric implications of the theory, as this indicates ways of performing equilibrium-based selection correction. As such, the model either brings existing work closer to empirical implementation, or enriches its theoretical foundations. Of course the model is affected by several shortcomings, which suggest directions for future research. Two of them are particularly important. The first is the model reliance on stringent parametric assumptions. Ideally, one wants to work with general distributions and identify or at least partially identify social interactions after considering non-random assignment to groups, at no further cost in terms of parametric assumptions. Brock and Durlauf (2004) suggest a number of possibilities in this direction, which have not been applied in empirical work yet. The second is the unobservability of the choice set at the first stage: in the data we have actual memberships, but not the alternatives one could have chosen. This means we need a theory of how people search and select memberships among alternative groups. This in turn can suggest instruments for the unobserved choice set.

Despite these shortcomings, the model can be applied to study, theoretically and in a tentative way empirically, several interesting cases that possibly involve stratification based on social interactions, such as parents spending resources to secure good associations to their kids in schools and neighborhoods, scholars trying to join stimulating departments, people joining clubs and social networks in general, employers selecting and stratifying the workforce, to mention a few. Many instances of social and economic inequality hinge upon the selection mechanism I have described, which also suggests equality and efficiency need not be traded off as one enlarges the perspectives of economic analysis. More importantly, this contribution is hopefully a useful stage in furthering an integrated study of economic and social phenomena.

9 Technical appendix

A1. Derivation of equations (15)-(16).

The argument here follows Ben-Akiva and Lerman (1985, pp. 287-288). The key fact is that if X_1 and X_2 are two independent extreme value (EV) random variables, with common scale parameter equal to 1, and means μ_1 and μ_2, then $\max(X_1, X_2)$ is EV distributed, with scale parameter equal to 1, and position parameter (corresponding to the mode) equal to $\log[\exp(\mu_1) + \exp(\mu_2)]$. Therefore, $\max_\omega \left(h_i\omega_i + J_g\omega_i m^e_{ig} + \varepsilon_{ig\omega}\right)$ is EV distributed, with position parameter equal to $\log \sum_{\omega} \exp \left(h_i \omega_i + J_g \omega_i m_{ig}^e \right)$ and scale parameter equal to 1. Since the expected value of an EV random variable is its mode plus Euler constant, γ, divided by the scale parameter, we have:

$$\begin{aligned} E \max_{\omega} \left(h_i \omega_i + J_g \omega_i m_{ig}^e + \varepsilon_{ig\omega} \right) &= \log \sum_{\omega} \exp \left(h_i \omega_i + J_g \omega_i m_{ig}^e \right) + \gamma \\ &= W_{ig} + \gamma . \end{aligned}$$

The probability that individual i chooses group g is:

$$\begin{aligned} & p_{ig} = \Pr(k_g - \rho_g + \varepsilon_{ig} + \max_{\omega} (h_i \omega_i + J_g \omega_i m_{ig}^e + \varepsilon_{ig\omega}) \\ \geq \ & k_{g'} - \rho_{g'} + \varepsilon_{ig'} + \max_{\omega} (h_i \omega_i + J_{g'} \omega_i m_{ig'}^e + \varepsilon_{ig'\omega})). \end{aligned}$$

A random variable is equal to its mean plus the stochastic part, therefore:

$$\max_{\omega} \left(h_i \omega_i + J_g \omega_i m_{ig}^e + \varepsilon_{ig\omega} \right) = W_{ig} + \gamma + \widetilde{\varepsilon}_{ig},$$

and p_{ig} can be rewritten as follows:

$$\begin{aligned} & p_{ig} = \Pr(k_g - \rho_g + W_{ig} + \varepsilon_{ig} + \widetilde{\varepsilon}_{ig} \\ \geq \ & k_{g'} - \rho_{g'} + W_{ig'} + \varepsilon_{ig'} + \widetilde{\varepsilon}_{ig'}). \end{aligned}$$

The purpose of the apparently obscure assumption that ε_{ig} is distributed such that the maximum of $V_i (g, \omega)$ is EV is to assure that $\varepsilon_{ig} + \widetilde{\varepsilon}_{ig}$ is EV distributed, with parameter β. This leads to equation (16).

A2. Proof of Proposition 1

Since the distribution of the random component of utility is continuous and differentiable, p_{hA}, $h = H, L$, is continuous and differentiable in ρ_A, and so is $f_H p_{HA} + f_L p_{LA}$. Equilibrium requires equality between the latter and α_A, the relative capacity of group A. Define $z_A (\rho_A, \rho_B) \equiv f_H p_{HA} + f_L p_{LA} - \alpha_A$ as the excess demand for membership in group A. This function is continuous, and furthermore, for any finite ρ_B

$$\begin{aligned}\lim_{\rho_A \to +\infty} z_A(\rho_A, \rho_B) &= -\alpha_A < 0 \\ \lim_{\rho_A \to -\infty} z_A(\rho_A, \rho_B) &= 1 - \alpha_A > 0\end{aligned}$$

By the intermediate value theorem, there exists a pair of prices (ρ_A^*, ρ_B^*) such that $z_A(\rho_A^*, \rho_B^*) = 0$. By Walras' law, the excess demand for membership in group B, $z_B(\rho_A, \rho_B)$, is also zero at (ρ_A^*, ρ_B^*). Therefore, the latter is a pair of prices clearing the memberships market. This proves existence. To prove uniqueness, one needs to show that the excess demand $z_g\left(\rho_g, \rho_{g'}\right)$, $g = A, B$, $g \neq g'$ is monotonic in ρ_g given $\rho_{g'}$, which implies it crosses the horizontal axis only once for any $\rho_{g'}$. This is in fact the case, since

$$\frac{\partial z_g\left(\rho_g, \rho_{g'}\right)}{\partial \rho_g} = \frac{\partial\left(f_H p_{Hg} + f_L p_{Lg}\right)}{\partial \rho_g} = \beta \sum_{h=H,L} f_h p_{hg}\left(p_{hg} - 1\right) < 0$$

Q.E.D.

A3. Proof of Proposition 3

We can consider either of the two groups and omit the group index, g. The right hand side of equation (30) represents an ordered set of functions, indexed by f_H. Such functions are convex combinations of the two functions $\tanh(H + Jm)$ and $\tanh(L + Jm)$. The tanh function is strictly increasing, and $H > L > 0$. This has two implications. First, these two functions are, respectively, the upper and lower bound of the set, i.e.

$$\tanh(H + Jm) > f_H \tanh(H + Jm) + (1 - f_H)\tanh(L + Jm) > \tanh(L + Jm),$$

for $f_H \in (0, 1)$. Second, for every $f'_H > f_H$,

$$\begin{aligned} & f'_H \tanh(H + Jm) + (1 - f'_H)\tanh(L + Jm) \\ > \; & f_H \tanh(H + Jm) + (1 - f_H)\tanh(L + Jm) \end{aligned} \tag{45}$$

i.e. the functions lie one above the other and do not cross, with "higher" functions associated with higher values of f_H. Based on the properties of the tanh function, propositions (U) and (M) can be rephrased in more detail as follows:

(U) *Equation* (30) *has a unique root* $m_1 > 0$*, such that the difference between the RHS and the LHS is positive for* $m < m_1$*, and negative for* $m > m_1$

(M) *Equation* (30) *has three roots,* $m_1 > 0$*,* $m_2 < 0$ *and* $m_3 < 0$*, such that the difference between the RHS and the LHS is positive for* $m < m_3$ *and* $m_2 < m < m_1$ *and negative for* $m > m_1$ *and* $m_3 < m < m_2$*.*

By assumption we have that

$$\begin{aligned} f_H &= 1 \Rightarrow \text{(U)} \\ f_H &= 0 \Rightarrow \text{(M)} \end{aligned}$$

Suppose there exist $f_H''' > f_H'' > f_H'$ such that

$$\begin{aligned} f_H &= f_H''' \Rightarrow \text{(M)} \\ f_H &= f_H'' \Rightarrow \text{(U)} \\ f_H &= f_H' \Rightarrow \text{(M)} \end{aligned}$$

Then there exists roots $m_3' < 0$ and $m_3''' < 0$ such that

$$\begin{aligned} & f_H'' \tanh\left(H + Jm_3'\right) + (1 - f_H'') \tanh\left(L + Jm_3'\right) - m_3' \\ > \; & f_H' \tanh\left(H + Jm_3'\right) + (1 - f_H') \tanh\left(L + Jm_3'\right) - m_3' = 0 \end{aligned}$$

and

$$\begin{aligned} & 0 = f_H''' \tanh\left(H + Jm_3'''\right) + (1 - f_H''') \tanh\left(L + Jm_3'''\right) - m_3''' \\ < \; & f_H'' \tanh\left(H + Jm_3'''\right) + (1 - f_H'') \tanh\left(L + Jm_3'''\right) - m_3''' \end{aligned}$$

The second inequality contradicts condition (45), since $f_H''' > f_H''$. Such contradiction implies that the unit interval can be partitioned into two disjoint subsets M and U, with $\mu < \upsilon$ for each $\mu \in M$ and $\upsilon \in U$, such that $f_H \in M$

implies (M) and $f_H \in U$ implies (U). Therefore, by the separating hyperplane theorem, there exists a unique value $\widetilde{f}_H$ such that

$$\begin{aligned} f_H &> \widetilde{f}_H \Rightarrow \text{(U)} \\ f_H &< \widetilde{f}_H \Rightarrow \text{(M)} \end{aligned}$$

Q.E.D.

References

[1] Bayer, Patrick, and Christopher Timmins (2005). "On the Equilibrium Properties of Locational Sorting Models." *Journal of Urban Economics*, 57, 462-77.

[2] Becker, Gary, and Kevin Murphy (2000). *Social Economics*, Cambridge Mass.: Harvard University Press.

[3] Ben-Akiva, Moshe (1973). *Structure of Passenger Travel Demand Models*. Ph.D. dissertation, Dept. of Civil Engineering, MIT, Cambridge Mass.

[4] Ben-Akiva, Moshe and Steven Lerman (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge Mass.: MIT Press.

[5] Bénabou, Roland (1993). "Workings of a City: Location, Education, and Production." *Quarterly Journal of Economics*, 108, 619-652.

[6] Bénabou, Roland (1996), "Equity and Efficiency in Human Capital Investment: the Local Connection." *Review of Economic Studies*, 63, 237-264.

[7] Blume, Lawrence, and Steven Durlauf (2003). "Equilibrium Concepts for Social Interaction Models." *International Game Theory Review*, 5, 193-209.

[8] Blume, Lawrence, and Steven Durlauf (2006). "Identifying Social Interactions: A Review." In *Methods in Social Epidemiology*, J. Oakes and J. Kaufman, eds., San Francisco: Jossey-Bass.

[9] Borsch-Supan, Axel (1990). "On the Compatibility of Nested Logit Models with Utility Maximization." *Journal of Econometrics*, 43, 373-388.

[10] Brock, William, and Steven Durlauf (2001a). "Discrete Choice with Social Interactions." *Review of Economic Studies*, 68, 235-260.

[11] Brock, William, and Steven Durlauf (2001b). "Interactions-Based Models." In *Handbook of Econometrics*, 5, J. Heckman and E. Leamer eds., Elsevier Science, North-Holland: 3297-3380.

[12] Brock, William, and Steven Durlauf (2004). "Identification of Binary Choice Models with Social Interactions." *Working Paper*, University of Wisconsin-Madison.

[13] Brock, William, and Steven Durlauf (2005). "Multinomial Choice with Social Interactions." In *The Economy as an Evolving Complex System III*, Lawrence Blume and Steven Durlauf (editors). Oxford University Press.

[14] Cooper, Russell, and Andrew John (1988). "Coordinating Coordination Failures in Keynesian Models." *Quarterly Journal of Economics*, 103, 441-464.

[15] de Bartolome, Charles (1990). "Equilibrium and Inefficiency in a Community Model with Peer Group Effects." *Journal of Political Economy*, 98, 110-133.

[16] Debreu, Gerard (1960). "Review of D. Luce, Individual Choice Behavior: A Theoretical Analysis." *American Economic Review*, 50, 186-188.

[17] Dubin, Jeffrey, and Douglas Rivers (1989/1990). "Selection Bias in Linear Regression, Logit and Probit Models." *Sociological Methods and Research*, 18, 360-390.

[18] Durlauf, Steven (1996). "A Theory of Persistent Income Inequality." *Journal of Economic Growth*, 1, 75-93.

[19] Durlauf, Steven (2004). "Neighborhood Effects." In *Handbook of Urban and Regional Economics*, 4, J. V. Henderson and J. F. Thisse eds., Elsevier Science, North-Holland, 2172-2142.

[20] Epple, Dennis, and Glenn Platt (1998). "Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Income." *Journal of Urban Economics*, 43, 23-51.

[21] Heckman, James (1979). "Sample Selection Bias as a Specification Error." *Econometrica*, 47, 153-161.

[22] Herriges, Joseph, and Catherine Kling (1996). "Testing the Consistency of Nested Logit Models with Utility Maximization." *Economic Letters*, 50, 33-39.

[23] Ichino, Andrea, and Armin Falk (2006). "Clean Evidence on Peer Pressure." *Journal of Labor Economics*, 24, 39-57

[24] Ioannides, Yannis, and Jeffrey Zabel (2004). "Interactions, Neighborhood Selection, and Housing Demand." *Working Paper*, Tufts University.

[25] Lee, Lung-Fei (1983). "Generalized Econometric Models with Selectivity." *Econometrica*, 51, 507-510.

[26] Manski, Charles (1993). "Identifications of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, 60, 531-542.

[27] Manski, Charles (2000). "Economic Analysis of Social Interactions." *Journal of Economic Perspectives*, 14, 115-136.

[28] McFadden, Daniel (1978). "Modelling the Choice of Residential Location." In *Spatial Interaction Theory and Planning Models*, A. Karlqvist, L. Lundqvist, F. Snickars and J. Weibull eds., North-Holland, 75-96.

[29] Milgrom, Paul and John Roberts (1990). "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities." *Econometrica*, 58, 1255-1277.

[30] Newey, Whitney, James Powell, and James Walker (1990). "Semiparametric Estimation of Selection Models: Some Empirical Results." *American Economic Review*, 80, 324-328.

[31] Sacerdote, Bruce (2001). "Peer Effects with Random Assignments: Results from Dartmouth Roommates." *Quarterly Journal of Economics*, 116, 681-704.

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

[33] Train, Kenneth (2003). *Discrete Choice Methods with Simulation*, Cambridge Mass.: MIT Press.

[34] Zimmerman, David (2003). "Peer Effects in Higher Education: Evidence from a Natural Experiment." *Review of Economics and Statistics*, 85, 9-23.