

Discrete Choice with Social Interactions and Endogenous Memberships

Giulio Zanella*

June 25, 2004

Abstract

This paper develops a theory of sorting and behavior, when the latter is subject to social influences. I extend the choice model with social interactions 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 structure 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.

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

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

*Department of Economics, University of Wisconsin-Madison, and University of Siena. E-mail: gzanella@wisc.edu.

Contents

1	Introduction	3
2	The model	7
3	Choice	12
3.1	Behavior	14
3.2	Group formation	14
3.3	Self-consistency	15
4	Equilibrium	17
4.1	Memberships market equilibrium	17
4.2	Behavioral equilibrium	19
4.3	Sorting equilibrium	21
5	Characterizing sorting equilibria	22
6	Welfare analysis	29
7	A numerical exercise	30
8	Econometrics	32
9	Discussion and conclusions	35
10	Technical appendix	40

ACKNOWLEDGMENTS. This paper is part of my doctoral dissertation, and has been 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 UW Madison 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, USA.

To choose a neighborhood is to choose neighbors

-Thomas Schelling, *Dynamic Models of Segregation*, 1971¹

1 Introduction

This paper describes a model of choice, under the assumption that choice is affected by social interactions, and individuals can select in advance their own social context. Following Brock and Durlauf (2003), by social interactions, or neighborhood effects, I mean direct interdependencies, not mediated by markets and enforceable contracts, between individual decisions and the decisions and characteristics of others within a group². Such interdependencies capture a variety of sociological and psychological effects. The paper tackles a theoretical issue, as well as an econometric one, although the former is instrumental to the latter. On the theoretical side, I am trying to organize a framework for merging two complementary, yet somehow separated, research programs on the aggregate consequences of social interactions. On one hand are models with endogenous memberships but exogenous neighborhood effects (e.g. de Bartolome, 1990, Bénabou, 1993 and 1996, Durlauf, 1996a,b). Such models introduce the notion of neighborhood effects in a Tiebout-type model: households sort into communities, and community affects offspring's human capital, or some other outcome, through an average community-level characteristic, such as income or human capital, that is *predetermined* (with respect to individual choice) by the sorting process itself. In this sense interactions are exogenous. On the other hand there are models with endogenous social effects but exogenous memberships, which focus on the microstructure of interactions, given group composition (e.g. Glaeser, Sacerdote and Scheinkman, 1996, Akerlof, 1997, Brock and Durlauf, 2001a and 2003). In

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

²Such interdependencies imply correlated behavior within groups, but the converse is not true. In a seminal paper, Manski (1993) devised three causes of such correlation: an *endogenous social effect* exists if one's behavior varies with the behavior of others within a group. An *exogenous* (or *contextual*) *social effect* exists if one's behavior varies with some predetermined characteristic of the group. Finally, a *correlated* (non social) *effect* is present if people behave similarly because they are similar in some respect. I stick to these definitions throughout the paper.

such models, individuals behavior is affected by *simultaneous* neighbors' behavior. In this sense interactions are endogenous. The development of a framework for the analysis of equilibrium group formation in presence of endogenous social interactions is a natural next step from a theoretical point of view. On the econometric side, the synthesis of such "orthogonal" perspectives is important because it indicates a way to solve two major empirical problems. Endogenous social effects have important implications for policy evaluation, because they imply a "social multiplier" of behavior. However, attempts to detect their presence are undermined by (1) an identification problem, and (2) self-selection bias³. Identification is a problem because it may be hard to distinguish the social effect from the effect of similar individual characteristics. Self-selection is a source of bias because purposeful decision makers will choose their memberships considering the costs and benefits associated with alternative groups. But if an individual has chosen her group, she has also chosen the value of the group-specific explanatory variables. When projecting individual behavior on such variables, the error term will contain other omitted variables that affect group choice, and so it will be correlated with the regressors: self-selection induces endogeneity. This is true for the social effect itself, since prospective social interactions may influence the preference for a group relative to others. For instance, think of a student who must decide whether to drop out high school or not. Suppose that the dropout decision is subject to an endogenous social effect, i.e. after controlling for individual and family characteristics, the student's dropout probability increases with the dropout rate in her classroom, or her residential neighborhood. Since parents choose a school and a residential neighborhood for the young person, if they are aware of the existence of social effects, they will choose memberships so as to minimize the dropout probability. Or suppose that we want to estimate an epidemic theory of welfare dependency, i.e. the effect of a high rate of welfare dependency in a certain group on the decision between work and welfare. A household may well be member of a certain group because it provides useful information about welfare eligibility, or because it reduces the stigma of being on welfare. In both examples, households are, *de facto*, choosing social interactions. Overlooking the possibility of group selection will lead to overestimated social effects. This is a version of the selection problem, whose study was pioneered by Heckman (1979): if

³These problems will be treated formally in Section 8. See also Brock and Durlauf (2001b).

individuals can sort themselves into the groups whose effect the researcher is trying to disentangle, then the researcher is working with a non-random sample. As one solves the endogeneity problem, e.g. using instrumental variables techniques, the social effect may vanish (see Evans, Oates, and Schwab, 1992). This is why considerations of self-selection into groups are unescapable in models with social interactions⁴. However, valid instruments may be hard to find. A popular way to solve the selection-problem, indeed a very fruitful one, is to use natural experiments (e.g. Sacerdote, 2001, Zimmerman, 2003) or controlled experiments (e.g. Ichino and Falk, 2003), in which people is exogenously assigned to the groups under study. This paper, following Brock and Durlauf (2001b, 2003) and Ioannides and Zabel (2002), presents a way to solve the selection problem out of the equilibrium properties of the sorting process, and so may be useful when good instruments are not readily available, and when randomization or extrapolations from it, are not possible. Such way consists of exploiting the equilibrium conditions produced by the aforementioned theoretical model, to build an econometric model corrected for self-selection. Since such selection correction will introduce nonlinearity, the identification problem discussed by Manski (1993) and Brock and Durlauf (2001b) is solved too. As I will show, such econometric model is actually useless if one of the main assumptions of the theory, i.e. that people rationally consider the existence of social interactions when choosing their memberships, is incorrect. Nonetheless, as I will argue, the model has the power to falsify the theory, and so, I think, it may be useful anyway. Beyond serving as a means to solve, in a microfounded way, the identification and the self-selection problems, the theoretical model is interesting in itself, and so I will devote considerable space to the analysis of its properties. The main theoretical implication of the model is that the sorting process and the neighborhood effect feedback reciprocally, and interact in characterizing the equilibrium distribution of memberships and behaviors. Positive complementarity between individual behaviors (in the sense of Cooper and John, 1988, and Milgrom and Roberts, 1990) is a stratifying force. Equilibrium stratification, along individual observable characteristics, is induced by the social effect, and is sustained by the equilibrium in the memberships market. The sorting process may generate an endogenous change in the qualitative

⁴Brock and Durlauf (2001b, 2003) discuss in-depth the issue of endogenous memberships as source of self-selection bias, and devise several ways to deal with it, including the approach I develop in this paper.

properties of the system, a phase transition, with the emergence of a “social trap”. In this case the outcome of the sorting process is inefficient. I will assume individual heterogeneity has an observable and an unobservable components. As a consequence, complete stratification along one dimension, a typical result in sorting models, is not a necessary condition for equilibrium. This echoes Epple and Platt (1998), who show that heterogeneity along both income and preferences prevents complete stratification in a Tiebout model. My framework is a possible behavioral underpinning for this kind of result. The main econometric implication of the model is that it allows to construct a Heckman correction term that solves both selection and identification problems that typically affect the “baseline” linear model with social effects. The first successful attempt in this direction is due to Ioannides and Zabel (2002). However, theirs is not an equilibrium model, and is mainly concerned with the estimation of neighborhood effects⁵. In this paper I make a similar attempt. The advantage of the procedure I develop is that it is derived from maximizing behavior, and equilibrium. The theoretical and the econometric issues are tackled using the same model: I extend the discrete choice model with social interactions developed by Brock and Durlauf (2001a,b and 2003) to allow for equilibrium group formation, when memberships can be bought on the market. Borrowing an idea in Brock and Durlauf (2003), I develop their logit model of behavior into a nested logit framework (Ben-Akiva 1973, McFadden, 1978), whose structure can nicely accommodate a model of sequential and interdependent choices of group and behavior. The paper is organized as follows. In section 2, I build the model in a very simple setting: two types of random-utility maximizer 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 neighborhood effects⁶. In section 3, using the random utility maximization hypothesis, I derive a nested logit model of probabilistic choice under self-consistent (i.e. ratio-

⁵An abstract theoretical model in which an action subject to interactions effects is chosen along with the group which exert that effect, also appears in Glaeser and Scheinkman (2001). However, their primary interest is the derivation of restrictions on the structure of variances as a possible route to identification.

⁶An appendix, available from the author upon request, generalizes this framework introducing multinomial behavior and more than two groups. However, the conclusions I can draw from the general model are much less clear-cut than those in the binary case, since as the dimension of the choice set grows, the relationship between the model parameters and the number of equilibria becomes very complex (see Brock and Durlauf, 2003, for a discussion).

nal) 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 determine the conditions under which the sorting process generates 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, there is more stratification when interactions are strong, and when the unobservable determinants of sorting have low dispersion. Integration is an equilibrium, but an unstable one. Section 6 formulates the planner's problem: I show that when a departure from integration leads to multiple behavioral equilibria in one group, the decentralized sorting process is inefficient. In section 7, I perform a numerical simulation to elucidate some of these results. Finally, in section 8, I turn to the econometric issues, and show how the model can help overcome identification and selection problems. Section 9 discusses some weak features of the model, and concludes.

2 The model

Imagine a population of I individuals. Each individual i , $i = 1, \dots, I$, is characterized by an observable real parameter, h_i , which belongs to the set $\Theta = \{H, L\}$, with $H > L$. The two types are distributed in the population according to the discrete distribution f_h , $h \in \Theta$. There are two groups, $G = \{A, B\}$, and each individual must be member of one group $g \in G$. The groups have the same fixed capacity, $I/2$, and are characterized by a set of *exogenous* variables, condensed into some index, k_g . Denote with f_{hg} the (discrete) distribution of h -types in group g . By definition, the distributions of types in the population and in the groups satisfy:

$$\begin{aligned}
 f_H + f_L &= 1 \\
 f_{Hg} + f_{Lg} &= 1 \quad g \in G
 \end{aligned}
 \tag{1}$$

Furthermore, for any type h , the number of h -type individuals in group A , which is equal to $(I/2) f_{hA}$, plus the number of individuals of the same type

in group B , $(I/2) f_{hB}$, must be equal to the number of h -type individuals in the population, $I f_h$. This implies that, for any type h :

$$f_{hA} + f_{hB} = 2f_h \quad (2)$$

The individual problem is to choose a group, $g \in \{A, B\}$, and a certain behavior. The latter is coded into a binary variable, $\omega \in \{-1, 1\}$. The price of membership in group g is ρ_g , which is determined in a perfectly competitive memberships market with fixed supply. Although group and behavior are chosen together by each individual, and simultaneously by all individuals, choice can be represented as a two-stage process. Such process is illustrated in figure 1. First a group is chosen. Once choices over memberships are mutually compatible, i.e. f_{hA} and f_{hB} turn out to be “equilibrium” distributions, individuals choose their behavior. I use the sequential representation to express the idea that individuals, when choosing their membership, will take into account its effect on the behavioral problem⁷. The set of terminal nodes in the decision tree is the choice set, $\Omega = \{A, B\} \times \{-1, 1\}$. This completes the skeleton of the model: it is abstract enough to encompass several possible situations of interest⁸.

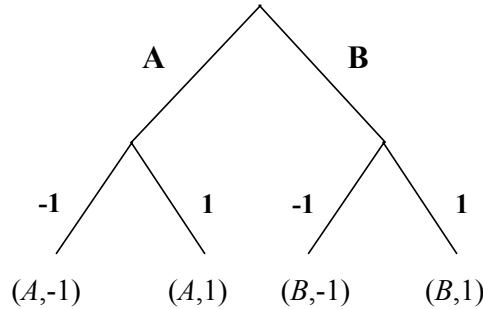


Figure 1. Individual decision tree: the final nodes are the four alternatives.

Through the rest of this section I will closely follow Brock and Durlauf (2001a), and I will extend their model to accommodate the choice structure

⁷The model can be easily extended to a genuinely two-stage problem.

⁸The “individuals” considered in the model can be persons, households, organizations, or governments. Therefore, groups can be anything that is pertinent to the identity of the individuals, e.g. classrooms, residential neighborhoods, industrial districts, trade agreements, or monetary unions.

depicted in figure 1. It is an extension in the sense that the choice set in their model is $\Omega = \{-1, 1\}$, i.e. individuals have no choice over memberships, and group composition is given. Using bold-faced letters for vectors, let's denote with $(\mathbf{g}, \boldsymbol{\omega})_{-i} \in \Omega^{I-1}$ the choice over groups and behaviors of individuals other than i . Assume that the utility of each individual is separable:

$$V_i(g_i, \omega_i) = v(g_i) + u_i(\omega_i, h_i) + S(g_i, \omega_i, \mu_i^e(\mathbf{g}, \boldsymbol{\omega})_{-i}) + \xi(g_i, \omega_i) \quad (3)$$

where (g_i, ω_i) denotes the choice of individual i . This function has four additive components: (1) $v(g)$ is a public payoff for all members of group g , regardless of behavior, and independent of type; (2) $u_i(\omega, h_i)$ is the private utility of behavior ω , regardless of membership, and is parameterized by i 's type; (3) $S(g, \omega, \mu_i^e(\mathbf{g}, \boldsymbol{\omega})_{-i})$ is a social utility term, and expresses the effects social interactions. It depends on one's choice, and on the joint probability measure $\mu_i^e(\mathbf{g}, \boldsymbol{\omega})_{-i}$, which expresses i 's subjective belief on the choice of other individuals; (4) $\xi(g, \omega)$ is a random utility term, which depends on both group and behavior. Agent i observes $\xi(g_i, \omega_i)$, but nobody else does. I also assume that $\xi(g_i, \omega_i)$ is identically and independently distributed across individuals, and that something (to be specified in a moment) about this distribution is common knowledge. In addition to separability, a second parametric assumption concerns the social utility term. I assume that individuals care only about mean behavior in their group, i.e. the only relevant moment of $\mu_i^e(\mathbf{g}, \boldsymbol{\omega})_{-i}$ is the subjective mean, denoted m_{ig}^e . Individuals are subject to a spillover from such mean behavior. The sign of the spillover depends on one's own behavior, and its weight in utility depends on a social interactions parameter, $J_g > 0$:⁹

$$S(g, \omega, \mu_i^e(\mathbf{g}, \boldsymbol{\omega})_{-i}) = J_g \omega m_{ig}^e \quad (4)$$

Assuming $I > 2$, the subjective mean, m_{ig}^e , can be represented as the average expected behavior of i 's "neighbors" in her own group, considering expected group composition:

$$m_{ig}^e = \left(\frac{I}{2} - 1\right)^{-1} \sum_{j \neq i} m_{ij|g}^e (\mu_i^e(g_j = g)) \quad (5)$$

⁹Such formulation of social utility, known as the "proportional spillovers" case, is equivalent, in the binary case, to a model with conformity effects (see Brock and Durlauf, 2001a).

On the right hand side of this expression, $m_{ij|g}^e(\cdot)$ is i 's subjective expectation of j 's behavioral choice, given that j is a member of group g . Such expectation is a function of the subjective marginal probability individual i places on j being a member of group g , $\mu_i^e(g_j = g)$. Notice that formulation (4) embodies endogenous social interactions, insomuch as behavioral choices are simultaneous¹⁰. I also assume that the public payoff from membership in group g depends only on the index of group characteristics and the membership price, in a linear way¹¹: $v(g) = k_g - \rho_g$. Furthermore, since ω has a binary support, we can rewrite the private component of utility, without loss of generality, as $u_i(\omega) = h_i\omega$. Indeed, the equality holds on the support of ω if we choose h_i such that $h_i = u(1)$ and $-h_i = u(-1)$, i.e. if $h_i = \frac{1}{2}(u_i(1) - u_i(-1))$.

Finally, let's turn to the random term. Since the alternatives in Ω may share unobserved attributes, the random term can be decomposed into 3 zero-mean random variables:

$$\xi(g_i, \omega_i) = \varepsilon_{ig\omega} + \varepsilon_{ig} + \varepsilon_{i\omega} \quad (6)$$

where the last two terms represent the portion of unobserved utility that varies, respectively, across groups and behaviors only. I am assuming 2 things about such random variables: (1) the share of the payoff due to unobserved elements varying across behaviors only is negligible, i.e.

$$\text{var}(\varepsilon_{i\omega}) = 0, \quad (7)$$

and (2) $\varepsilon_{ig\omega}$ and ε_{ig} are independent across individuals, groups, and behaviors. Such assumptions, of course, are not innocuous, and so deserve a word of comment. The first one has a socioeconomic meaning: when $\text{var}(\varepsilon_{i\omega}) = 0$ all the unobserved utility from behavioral alternatives is reflected in $\varepsilon_{ig\omega}$, i.e. depends on the social context. The assumption of orthogonality is more questionable in the part where it requires $\varepsilon_{ig\omega} \perp \varepsilon_{ig}$ for all i . This implies that an individual's unobservable preference for a particular behavior at the second stage of choice (given membership) is independent of her unobserv-

¹⁰Such interactions exhibit a strategic complementarity aspect: if the majority in the group is choosing $\omega = 1$, i.e. m_{ig}^e is positive, then i has an incentive to choose $\omega = 1$ too. Why it should be so? Usually, sociological and psychological arguments are used to justify the incorporation of social effects into utility functions. Samuelson (2004) has recently proposed an evolutionary argument, in the context of the theory of relative consumption.

¹¹Linearity helps to isolate the role of prices in the model.

able preference for the group she chose at the first stage. This can be very implausible in many contexts. Therefore, such limitation of the model must be kept in mind. An important implication of the assumptions on $\xi(g, \omega)$ concerns the covariance of utilities across groups. For any behavior ω , we have

$$\begin{aligned} \text{cov}(V_i(g, \omega), V_i(g', \omega)) &= \text{cov}(\varepsilon_{ig\omega} + \varepsilon_{ig}, \varepsilon_{ig'\omega} + \varepsilon_{ig'}) \\ &= E((\varepsilon_{ig\omega} + \varepsilon_{ig})(\varepsilon_{ig'\omega} + \varepsilon_{ig'})) = 0 \end{aligned}$$

where the last equality follows from the independence assumption. Therefore, the utilities of any particular behavior are independent across groups. This property is known as *independence from irrelevant nests* (IIN). Such terminology, which will be used through the rest of the paper, is due to the fact that the two main branches of the decision tree depicted in figure 1 resemble “nests”, and consequently the alternatives in the choice set are “nested”. The IIN property replaces the better known *independence from irrelevant alternatives* (IIA), which would beset the model if the alternatives in Ω were not nested, as instead they are (figure 1). Therefore, the representation of choice as a two-stage process makes the model more accurate¹².

Finally, in order to simplify the analysis and obtain a closed-form solution, I assume that $\varepsilon_{ig\omega}$ is extreme value (EV) distributed¹³, with scale parameter normalized to 1, and position parameter γ ,

$$F_{\varepsilon_{ig\omega}}(x) = \exp(-\exp(-x - \gamma)), \quad (8)$$

and that ε_{ig} is distributed such that the maximum of utility with respect to behavior is EV distributed with scale parameter β_γ and position parameter γ :

¹²As is well known, the IIA property signals models with independent disturbances. Such a property implies that the ratio between the probabilities of choosing any two alternatives is independent of the characteristics of all other alternatives. The property generates choice paradoxes such as the red bus/blue bus one (Debreu, 1960), and so is not a desirable property in several contexts.

¹³The extreme value distribution, also known as double-Weibull, double-exponential (for obvious reasons), and Gumbel distribution (after E. J. Gumbel originated it in the 1940s as a tool for predicting floods), is characterized by two parameters (η, β) setting, respectively, the position and the scale of the distribution. Under the normalization $\eta = -\frac{\gamma}{\beta}$, where $\gamma \simeq .577$ is the Euler constant, its mean and variance are, respectively, zero and $\frac{\pi^2}{6\beta^2}$.

$$F_{\varepsilon_{ig}} : F_{\varepsilon_{ig}} \left(\max_{\omega} V_i(g, \omega) \right) = \exp \left(- \exp \left(-\beta_{\gamma} \max_{\omega} V_i(g, \omega) - \beta_{\gamma} \gamma \right) \right) \quad (9)$$

where $\gamma \simeq .577$ is the Euler constant¹⁴. One can show that β_{γ} is a measure of the correlation between the utilities of different behaviors in a given group: the higher β_{γ} the lower the correlation¹⁵. This concludes the description of the model. Utility, omitting the individual index in the chosen alternative to lighten the notation, can be rewritten as:

$$V_i(g, \omega) = k_g - \rho_g + h_i \omega + J_g \omega m_{ig}^e + \varepsilon_{ig\omega} + \varepsilon_{ig} \quad (10)$$

Equations (6) through (9) describe what is known as the “nested logit error”. Therefore, we are ready to generate choice probabilities from a nested multinomial logit model (NMNL)¹⁶.

3 Choice

Each individual maximizes utility, i.e. solves:

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

Given the meaning of the random term in the model, the solution to this problem, from the viewpoint of individuals other than i , is a joint probability

¹⁴The conventional reason for this parametrization is that the difference between two identically EV distributed random variables is logistically distributed. Since the logistic distribution approximates the normal distribution, our parametric assumptions can be justified on the basis of the central limit theorem: the difference between the unobservable components of the utilities of any pair of alternatives is due to many unobservable factors, whatever their distribution may be.

¹⁵Therefore, the assumption that β_{γ} is constant across groups is not always appropriate (see Train, 2003, p.85).

¹⁶Alternatively (McFadden, 1978) the nested logit model can be derived assuming that $\xi_i(\omega, g)$ is distributed according to some Generalized EV distribution, and independently across groups, such that each marginal distribution is EV. The approach I follow (discussed in detail in Ben-Akiva and Lerman, 1985, Anderson, DePalma and Thisse, 1992, and Train, 2003) allows us to express directly the choice probabilities as the product of a marginal and a conditional logistic probabilities, which is more useful for theoretical purposes. The two formulations are, of course, equivalent.

distribution over (g_i, ω_i) . That is to say, while individuals choose deterministically a group and a behavior, each individual, ex-ante, can only cast a probability distribution over others' choice. Such distribution is then used to form an expectation about mean behavior in the two groups. The solution to problem (11) can be represented as follows. At the first stage, each individual maximizes utility with respect to membership, given that optimal behavior, denoted ω^* , will be chosen at the second stage, and given beliefs on other individuals' choices over groups, described by the probability measure $\mu_i^e(\mathbf{g}_{-i})$:

$$\begin{aligned} g_i &= \arg \max_g V_i(g, \omega^*) \\ &= \arg \max_g \left[k_g - \rho_g + \varepsilon_{ig} + \max_{\omega} (h_i \omega + J_g \omega m_{ig}^e + \varepsilon_{ig\omega}) \right] \end{aligned} \quad (12)$$

At the second stage, each individual maximizes utility with respect to behavior, given memberships, and given beliefs on mean behavior in the chosen group, g_i , described by the conditional probability measure $\mu_i^e(\omega_{-i}|\mathbf{g}_{-i})$:

$$\omega_i = \arg \max_{\omega} V_i(g_i, \omega) \quad (13)$$

Exploiting the foregoing parametric assumptions, such choice rules allow the derivation of *objective* choice probabilities, proceeding from the lower level of the tree¹⁷.

¹⁷The nested logit model is affected by a complication: choice rules (12)-(13) define a consistent random utility maximization problem, provided that the assumption of a nested logit structure, equations (6) through (9), is such that the density of $\varepsilon_{ig\omega} + \varepsilon_{ig}$ is non negative. The complication arises from the fact that the sum of two EV random variables is not itself EV distributed. A sufficient condition for consistency with random utility maximization is $\beta_{\gamma} \leq 1$. This condition is known as the Daly-Zachary-McFadden condition, but has been shown to be unnecessarily strong by Borsch-Supan (1990). A simple necessary and sufficient condition for compatibility with the random utility maximization hypothesis when there are two stages of choice and each nest contains less than four alternatives, has been devised by Herriges and Kling (1996). The condition involves a set of simple relations between the choice probabilities at the first stage and β_{γ} . Herriges and Kling's (1996) result corrects a previous theorem due to Borsch-Supan (1990), the first to extend the Daly-Zachary-McFadden condition.

3.1 Behavior

Rule (13) implies that the probability that individual i chooses ω , given membership in group g , is:

$$\begin{aligned} p_{i\omega|g} &= \Pr(V_i(g, \omega) > V_i(g, -\omega)) \\ &= \Pr(\varepsilon_{ig-\omega} - \varepsilon_{ig\omega} \leq 2h_i\omega + 2J_g\omega m_{ig}^e) \end{aligned} \quad (14)$$

From the properties of the EV distribution we know that the difference between $\varepsilon_{ig-\omega}$ and $\varepsilon_{ig\omega}$ is logistically distributed:

$$F_{\varepsilon_{ig-\omega} - \varepsilon_{ig\omega}}(x) = \frac{1}{1 + \exp(-x)} \quad (15)$$

Using equations (14) and (15), it is straightforward to derive the following logit choice probability for individual behavior, conditional on membership:

$$p_{i\omega|g} = \frac{\exp(h_i\omega + J_g\omega m_{ig}^e)}{\exp(h_i + J_g m_{ig}^e) + \exp(-h_i - J_g m_{ig}^e)} \quad (16)$$

This is the binary choice model with social interactions studied by Brock and Durlauf (2001a,b).

3.2 Group formation

The remaining part of the problem, involving rule (12), is simplified by the fact that the maximum of an EV random variable is itself EV. It can be shown (see technical appendix) that the expected maximum utility with respect to behavior has the following convenient expression:

$$E \max_{\omega} (h_i\omega + J_g\omega m_{ig}^e + \varepsilon_{ig\omega}) = \log \sum_{\omega=-1,1} \exp(h_i\omega + J_g\omega m_{ig}^e) + \gamma \quad (17)$$

where γ is Euler constant.

Therefore, expected maximum utility with respect to ω is simply the natural log of the denominator of the logit probability (16), plus a constant.

Following Ben-Akiva (1973), the quantity $W_{ig}^e \equiv \log \sum_{\omega} \exp(h_i \omega + J_g \omega m_{ig}^e)$ can be regarded, up to the constant γ , as the appropriate measure of the expected value of group g for individual i . This quantity is known as *inclusive utility*¹⁸, and represents the value of the behavioral “choice situation” in a group. Rule (12) implies the following choice probability over groups (see technical appendix):

$$p_{ig} = \frac{\exp(\beta_{\gamma}(k_g - \rho_g + W_{ig}^e))}{\exp(\beta_{\gamma}(k_A - \rho_A + W_{iA}^e)) + \exp(\beta_{\gamma}(k_B - \rho_B + W_{iB}^e))} \quad (18)$$

Equations (16) and (18) are linked by the inclusive utility of group g for individual i . The meaning of this link is that the inclusive utility of a certain 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, subject to the social effect, is chosen. This point is the key to the whole model, since expected social interactions turn out to be a crucial determinant of group composition, and of actual behaviors.

3.3 Self-consistency

To close the model, we need to specify the subjective belief terms which appear in equation (5). Such beliefs are crucial, since they determine expected mean behavior, i.e. the medium of social interactions. We have just seen that the model produces choice probabilities over groups and behaviors, therefore rational agents must form their expectations using these probabilities and the assumption that the other agents do the same. In other words, I am assuming rational expectations as an easy way to close the model. Therefore, the subjective beliefs $m_{ij|g}^e$ and $\mu_i^e(g_j = g)$ must be equal to the objective values generated by the model. As for membership, this means that the probability individual i attributes to j being a member of group g must be equal to the objective membership probability, as generated by equation (18):

$$\mu_i^e(g_j = g) = p_{jg} \quad \forall i, j \quad (19)$$

¹⁸A.k.a. inclusive value, log-sum term, and accessibility measure. In logistic models of consumer demand, the equivalent of the quantity $\beta_{\gamma} E[\max_{\omega} V_i(g, \omega)]$ is expected consumer surplus.

As for behavior, expected individuals behaviors must be computed using a mathematical expectation and the rules of probability:

$$\begin{aligned} m_{ij|g}^e(\mu_i^e(g_j = g)) &= m_{ij|g}^e \mu_i^e(g_j = g) \\ &= E(\omega_j|g) p_{jg} = (p_{j1|g} - p_{j-1|g}) p_{jg} \quad \forall i, j \end{aligned} \quad (20)$$

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

$$m_{ig} = \left(\frac{I}{2} - 1\right)^{-1} \sum_{j \neq i} E(\omega_j|g) p_{jg} \quad (21)$$

i.e. the self consistent version of equation (5). When the population is large enough, this expectation is well approximated by the average of individual expected behaviors in the group¹⁹, m_g :

$$m_g = \left(\frac{I}{2}\right)^{-1} \sum_i E(\omega_i|g) p_{ig} \quad (22)$$

Since I have imposed rational expectations, and since we are working with choice probabilities, all heterogeneity in the population is now due to the individual type, $h_i \in \{H, L\}$. Therefore, we can replace the individual index, i , with the type index, h , and rewrite (22) as follows:

$$m_g = 2 \sum_{h \in \Theta} f_h E(\omega_h|g) p_{hg} \quad (23)$$

The right hand side of equation (23) will turn out to have a very convenient expression once the memberships market equilibrium condition is imposed. For the moment, substituting m_g into the expression of W_{hg}^e , the probability distribution of memberships can be rewritten in objective form:

$$p_{hg} = \frac{\exp(\beta_\gamma(k_g - \rho_g + W_{hg}))}{\sum_{\nu=A,B} \exp(\beta_\gamma(k_\nu - \rho_\nu + W_{h\nu}))} \quad (24)$$

¹⁹The difference is due to the heterogeneity in h , since each individual “takes out” her own h when computing m_{ig} .

where

$$W_{hg} = \log \sum_{\omega} \exp(h\omega + J_g \omega m_g) \quad (25)$$

4 Equilibrium

An equilibrium is defined as a set of prices, $\{\rho_A^*, \rho_B^*\}$ labeled *memberships market equilibrium*, a set of expected mean behaviors, $\{m_A^*, m_B^*\}$ labeled *behavioral equilibrium*, and a set of membership probabilities, $\{p_{hA}^*, p_{hB}^*\}_{h \in \Theta}$ labeled *sorting equilibrium*, such that, for each individual, utility is maximized according to rules (12)-(13), and the memberships market clears. Let's analyze in detail each of the three components of the equilibrium.

4.1 Memberships market equilibrium

The sum of the individual choice probabilities over group g , given by equation (24), is equal to the expected number of individuals choosing g . This is the ex-ante aggregate demand for membership in group g . On the other side of the market, the supply of memberships in each group is fixed, and equal to $I/2$. Therefore, market clearing requires:

$$\sum_{h \in \Theta} f_h p_{hg} = \frac{1}{2} \quad (26)$$

We can conclude that a unique memberships market equilibrium exists:

Proposition 1 *A unique set of membership prices, $\{\rho_A^*, \rho_B^*\}$, solving equation (26) exists, for any pair of expected mean behaviors, $\{m_A, m_B\}$.*

Proof. See Technical Appendix.

The number of h -type individuals demanding membership in group g is $I f_h p_{hg}$. Since the size of each group is $I/2$, their fraction is $2 f_h p_{hg}$. When equilibrium condition (26) holds, such fraction, for $h \in \Theta$, is the distribution of types in group g . Therefore, in equilibrium:

$$f_{hg} = 2f_h p_{hg} \quad (27)$$

Notice that the membership probability p_{hg} , and the discrete density f_{hg} are different objects²⁰. An important implication of the memberships market equilibrium is a relation between membership prices and composition of groups. To see this, notice that the difference between equilibrium prices²¹ can be recovered using equilibrium condition (26) and choice probabilities (24). The algebraic expression of such difference, in general, is cumbersome. However, in the special case $f_H = f_L = \frac{1}{2}$, i.e. the types are equally represented in the population, such expression reduces to

$$\rho_A^* - \rho_B^* = (k_A - k_B) + \frac{1}{2}\beta_\gamma [(W_{HA} + W_{LA}) - (W_{HB} + W_{LB})] \quad (28)$$

The portion of the price differential that is not due to differences between exogenous group characteristics, i.e. $k_A - k_B$, reflects the different values of social interactions in the groups, due to different group composition. Therefore, such portion can be properly labeled “social premium”, s :

$$s = \frac{1}{2}\beta_\gamma [(W_{HA} + W_{LA}) - (W_{HB} + W_{LB})] \quad (29)$$

In the special case under analysis, the social premium reduces to a function of the difference between aggregate inclusive utilities in the two groups. Since inclusive utility carries information on expected social interactions, a group commands a positive social premium if it offers, on average across types, the “best interactions” (from an individual’s viewpoint), which are associated with the highest value of the choice situation for the two types. Such highest value is $\max\{(W_{HA} + W_{LA}), (W_{HB} + W_{LB})\}$. In the general case, $f_H \neq f_L$, the social premium reflects population composition as well. Given the existence of a social premium, a natural question is: does the memberships market eliminate the externalities generated by social interactions? The answer is that it eliminates the first stage externalities due to migration (since prices reflect inclusive utilities), but not the second stage externality stemming from the social component of utility.

²⁰They assume the same value in the special case in which the two types are equidistributed in the population, i.e. when $f_H = f_L = \frac{1}{2}$.

²¹Only the difference matters for memberships probabilities.

4.2 Behavioral equilibrium

Substituting expression (27) into (23), expected mean behavior in group g can be conveniently expressed as²²

$$\begin{aligned} m_g &= \sum_{h \in \Theta} f_{hg} E(\omega_h | g) \\ &= \sum_{h \in \Theta} f_{hg} \tanh(h + J_g m_g) \end{aligned} \quad (30)$$

where the second equality derives from the fact that $E(\omega_h | g) = p_{h1|g} - p_{h-1|g}$, and from the definition, for any polynomial P , $\tanh(P) \equiv (e^P - e^{-P}) / (e^P + e^{-P})$. A behavioral equilibrium is a fixed point of such equation, i.e. a value m_g^* in each group solving:

$$m_g^* = f_{Hg} \tanh(H + J_g m_g^*) + (1 - f_{Hg}) \tanh(L + J_g m_g^*) \quad (31)$$

We can conclude that a behavioral equilibrium exists:

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

Proof. *Since $m_g : [-1, 1] \rightarrow [-1, 1]$, we can apply Brouwer's fixed point theorem and conclude that one or more m_g^* exist. Q.E.D.*

From Brock and Durlauf (2001a), we know that: (i) there exist either one or three such equilibrium mean behaviors in each group (figure 2), depending on the interplay between the individual type and the magnitude of the social effect; (ii) in case of multiplicity, the equilibria are Pareto-rankable; (iii) the intermediate equilibrium is unstable if individuals can revise their behavior.

²²The second equality derives from the useful expression for the conditional expectation in the binary model, $E(\omega | g) = p_{i1|g} - p_{i-1|g}$, and using the definition $\tanh(P) = \frac{e^P - e^{-P}}{e^P + e^{-P}}$ for any polynomial P .

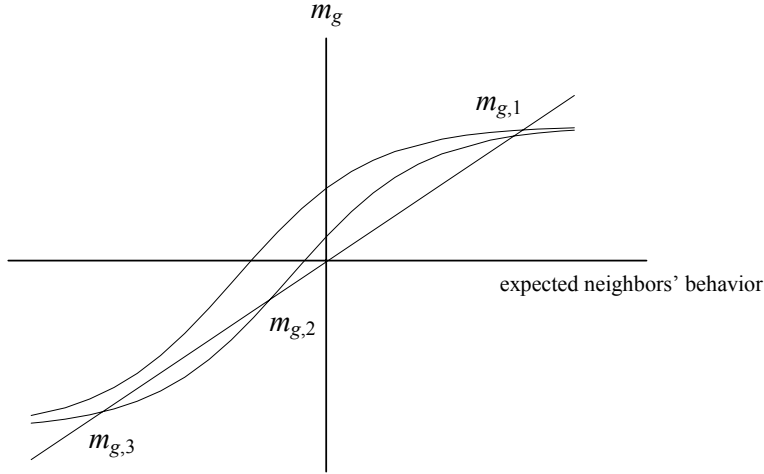


Figure 2. Unique and multiple behavioral equilibria.

A precise relation exists between multiple equilibria and the process of sorting. Such relation is an extension of Brock and Durlauf's (2001a) results, made possible by the endogenization of memberships, and can be stated formally as follows:

Proposition 3 *If $J_g > 1$, then there exists a threshold $\tilde{h}_g(f_{Hg}, J_g)$, which is a function of the proportion of H-types and of the magnitude of interactions in group g , such that:*

- (a) *if $f_{Hg}H + (1 - f_{Hg})L > \tilde{h}_g$, then equation (31) has a unique root with positive sign, $m_{g,1}^*$.*
- (b) *if $f_{Hg}H + (1 - f_{Hg})L < \tilde{h}_g$, then equation (31) has three roots, one with positive sign, $m_{g,1}^*$, and two with negative sign, $m_{g,2}^*$ and $m_{g,3}^*$.*

Proof. See Technical Appendix.

An important consequence is the following.

Corollary. *If $H > \tilde{h}_g > L$, then there exists a critical fraction of H-types in group g , $\tilde{f}_{Hg}(J_g)$, which is a function of the magnitude of interactions, above which group g has a unique behavioral equilibrium, and below which it experiences a phase transition, in the form of emergence of multiple behavioral equilibria.*

Proof. See Technical Appendix.

In words, proposition 3 and its corollary assert that multiple equilibria appear in group g when the average type falls below a critical threshold. This phenomenon is analogous to a physical phase transition, and it is more likely to occur if L is sufficiently small to generate multiple equilibria in a hypothetical group made up of L -types only, and if H is sufficiently large. Since both the average h and the threshold are endogenously determined by the sorting process, a phase transition may occur endogenously if the H -types have, for instance, a tendency to concentrate in one group. Section 5 is devoted to this issue.

4.3 Sorting equilibrium

The third component of equilibrium, namely sorting equilibrium, is trickier than the previous ones. I will simplify the analysis, at the cost of two assumptions. Denote with $\mu(m_{g,z}^*)$ the common belief over behavioral equilibrium z , $z = 1, 2, 3$, in group g . Assume (1) that the probability distribution over behavioral equilibria in the two groups is common knowledge before sorting takes place, i.e. individuals know $\mu(m_{g,z}^*)$, $z = 1, 2, 3$. This means they have a fairly accurate probabilistic idea of what the two groups will look like in terms of behavioral choices. And assume (2) that $\mu(m_{g,2}^*) = 0$, because of the instability of $m_{g,2}^*$. This means individuals believe the system will not remain for a long time at the unstable equilibrium. Under such assumptions, equilibrium inclusive utility can be rewritten as an average of the equilibria at the second stage:

$$W_{hg}^* = \sum_z \mu(m_{g,z}^*) \log [\exp(h + J_g m_{g,z}^*) + \exp(-h - J_g m_{g,z}^*)] \quad (32)$$

Since each behavioral equilibrium solves equation (31), $m_{g,z}^*$ is a function of f_{Hg} , and so, from (27), of p_{Hg} :

$$m_{g,z}^* = m_z^*(p_{Hg}) \quad z = 1, 2, 3 \quad (33)$$

Therefore, substituting equation (33) into (32), we can rewrite equation (24) as:

$$p_{hg} = \frac{\exp(\beta_\gamma (k_g - \rho_g^* + W_{hg}^*(p_{Hg})))}{\sum_{\nu=A,B} \exp(\beta_\gamma (k_\nu - \rho_\nu^* + W_{h\nu}^*(p_{H\nu})))} \quad (34)$$

This is the equilibrium probability function over individual membership. It is now straightforward to prove that a sorting equilibrium exists.

Proposition 4 *A set of equilibrium choice probabilities over groups, $\{p_{hA}^*, p_{hB}^*\}$ exist for each type h .*

Proof. *Since $p_{Hg} = 1 - p_{Hg'}$, $g \neq g'$, equilibrium inclusive utility (32) is a function of p_{Hg} only. So, $p_g : [0, 1] \rightarrow [0, 1]$, and we can apply Brouwer's fixed point theorem to conclude that one or more p_{ig}^* exist. Q.E.D.*

Given any distribution of types in the population, a sorting equilibrium determines an equilibrium distribution of types across groups, $\{f_{hA}^*, f_{hB}^*\}$, $h \in \Theta$, according to equation (27). The following are natural questions to ask in a model with sorting: (1) is the sorting equilibrium unique? (2) Is there a tendency towards stratification along the individual type? (3) Is the decentralized sorting equilibrium efficient? The next section addresses the first two questions, showing that the answers are respectively “no” and “yes”. Section 6 addresses the third, showing that the answer depends on the parameters. Of course, the answers to such questions are interwound with the behavioral side of the model, and notably with social interactions.. In the meanwhile, propositions 1, 2 and 4 allow to establish existence of a macro equilibrium:

Proposition 5 *An aggregate equilibrium, i.e. a triple $\{\rho_A^*, \rho_B^*\}$, $\{p_{hA}^*, p_{hB}^*\}_{h \in \Theta}$, and $\{m_A^*, m_B^*\}$, exists in the binary choice model with two endogenous groups.*

5 Characterizing sorting equilibria

Let's start with question (2), i.e. the stratification issue. A simple stratification index, ς , is the maximum distance, with respect to individual type, between membership probabilities in the two groups:

$$\varsigma = \max_h |p_{hA} - p_{hB}| \quad (35)$$

By construction, such index is bounded: $0 \leq \varsigma < 1$, with $\varsigma \simeq 1$ being a state of segregation. So, for instance, if in equilibrium group A is chosen with probability 0.8 by H -types, and with probability 0.4 by L -types, then the degree of equilibrium stratification is $\varsigma = \max\{0.6, 0.2\} = 0.6$.²³ A convenient way to answer the stratification question, is to define *isoprobability curves* in the $m_g - \rho_g$ space. This is a natural extension of bid-rent analysis²⁴ to a probabilistic model. An isoprobability curve represents the combination of membership price and expected mean behavior such that membership probability is constant. “Higher” curves in the $m_g - \rho_g$ space are associated with a lower probability level. The slope at a point (m_g, ρ_g) is the marginal rate of substitution between membership price and mean behavior at that point, for type h . The rate has a simple expression in the binary case:

$$MRS_{m\rho}^h = -\frac{\partial p_{hg}/\partial m_g}{\partial p_{hg}/\partial \rho_g} = J_g \tanh(h + J_g m_g) \quad (36)$$

Such ratio is increasing in h , therefore the single crossing property holds for isoprobability curves of different types. The property is illustrated in figure 3, which illustrates two possible isoprobability curves of different types.. Notice that the intersection of two curves at equilibrium probability levels is necessarily an aggregate equilibrium. An increasing ratio 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, m_g . This is a necessary condition for the degree of stratification to increase with m_g . However, because of the unobservable preference for memberships, the condition is not sufficient²⁵. The marginal rate of substitution reveals that the slope of isoprobability curves increases in the magnitude of social interactions, J_g : the

²³Using equation (26), it is easy to check that these values are part of a macro equilibrium if $f_H = \frac{1}{4}$ and $f_L = \frac{3}{4}$.

²⁴Bid-rent analysis is the traditional approach to stratification across locations. See Epple (2003) for a synthetic overview.

²⁵In Bénabou’s (1996) model, a similar single crossing condition is sufficient for full stratification of households according to their human capital, since the model is deterministic. Furthermore, he interprets the membership price as a rent, and assumes the H -types can always outbid the L -types, which leads automatically to an exchange of memberships. However, if one complicates the model, for instance assuming that membership means

more important interactions are, the more individuals are willing to pay for membership in a group with higher expected mean behavior. So, because of single crossing, a larger interactions parameter is associated with stronger incentives to segregate.

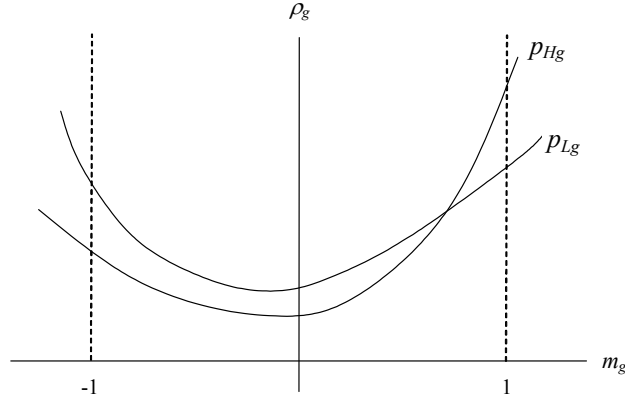


Figure 3. Isoprobability curves of the two types.

Let's now tackle question (1): how many sorting equilibria exist? In the rest of this section I will prove the following result:

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 ($\zeta = 0$), and is unstable. The other two are symmetric, are characterized by a certain degree of stratification ($\zeta > 0$), and are stable.*

To prove that integration is always an equilibrium, notice that when the memberships market clears, $\zeta = 0$ is equivalent to $p_{hA} = p_{hB}$, for both types. Hence, using equations (2) and (27):

$$f_{hA} = f_{hB} = f_h \tag{37}$$

homeownership, it may be hard to justify why a L -types household is willing to sell its house now, rather than keeping an asset whose value is increasing because of higher demand.

for both types. In other words, integration implies that group composition reflects the proportion of types in the population. Hence:

$$p_{hg} = \frac{1}{2} \quad h \in \Theta, \quad (38)$$

since the two groups have the same composition. When (37) and (38) hold, expected mean behavior is the same in the two groups, and so are the inclusive utilities, for both types: $W_{hA} = W_{hB}$. In this case, of course, no group can command a social premium. Indeed, solving $p_{hA} = p_{hB}$ from equation (34), and imposing $W_{hA} = W_{hB}$, it turns out that

$$\rho_A - \rho_B = k_A - k_B \quad (39)$$

i.e. the membership price difference reflects only exogenous group characteristics. Using these results, and rewriting (34) as

$$p_{hg} = \frac{1}{1 + \exp(\beta_\gamma (k_{g'} - k_g + \rho_g - \rho_{g'} + W_{hg'}^*(p_{hg}) - W_{hg}^*(p_{hg})))}, \quad (40)$$

where $g \neq g'$, it's straightforward to show that $p_{hg} = \frac{1}{2}$ solves equation (40). Therefore it is a fixed point, i.e. integration is a sorting equilibrium. However, it is unstable. To see this, it is sufficient to perform the following thought experiment. Suppose that the system is at the integrated sorting equilibrium, and perturb it, increasing m_A beyond its equilibrium level. As a consequence, the inclusive utility of group A , and membership probabilities in A , increase for both types. This will push up the membership price in A , because the supply of memberships is rigid. Higher prices will exert a downward pressure on membership probabilities, counterbalancing the initial increase. However, because of single-crossing, the downward pressure on p_{LA} is higher than that on p_{HA} . Therefore, individuals will revise upward both f_{HA} and f_{LB} , according to equation (27). The following lemma guarantees that expected mean behavior will not revert to the initial equilibrium:

Lemma 1. *If $H > L$, and m_A^* and m_B^* are stable roots of (31), then m_A increases (decreases), and m_B decreases (increases) as f_{HA} increases (decreases).*

Proof. See Technical Appendix.

As a consequence, since I have assumed $\mu(m_{g,2}^*) = 0$, inclusive utilities of the two types in the two groups must diverge. Membership prices will diverge too, and a positive social premium appears. The social premium brings

the memberships market to its new equilibrium. The left panel of figure 4 illustrates the adjustment mechanism²⁶. A new equilibrium is necessarily reached, since m_A is bounded, and cannot rise indefinitely. At the new equilibrium, there is a positive degree of stratification. The adjustment process in group B , illustrated in the left panel of figure 4, is symmetric, since group capacity is fixed. But there is an important difference: multiple equilibria may appear, for certain parameters values. In this case, ignoring the unstable behavioral equilibrium, there are two sets of isoprobability curves which are compatible with individual rationality. As a consequence, in group B the membership price falls to a level between the levels implied by these different sets of curves.

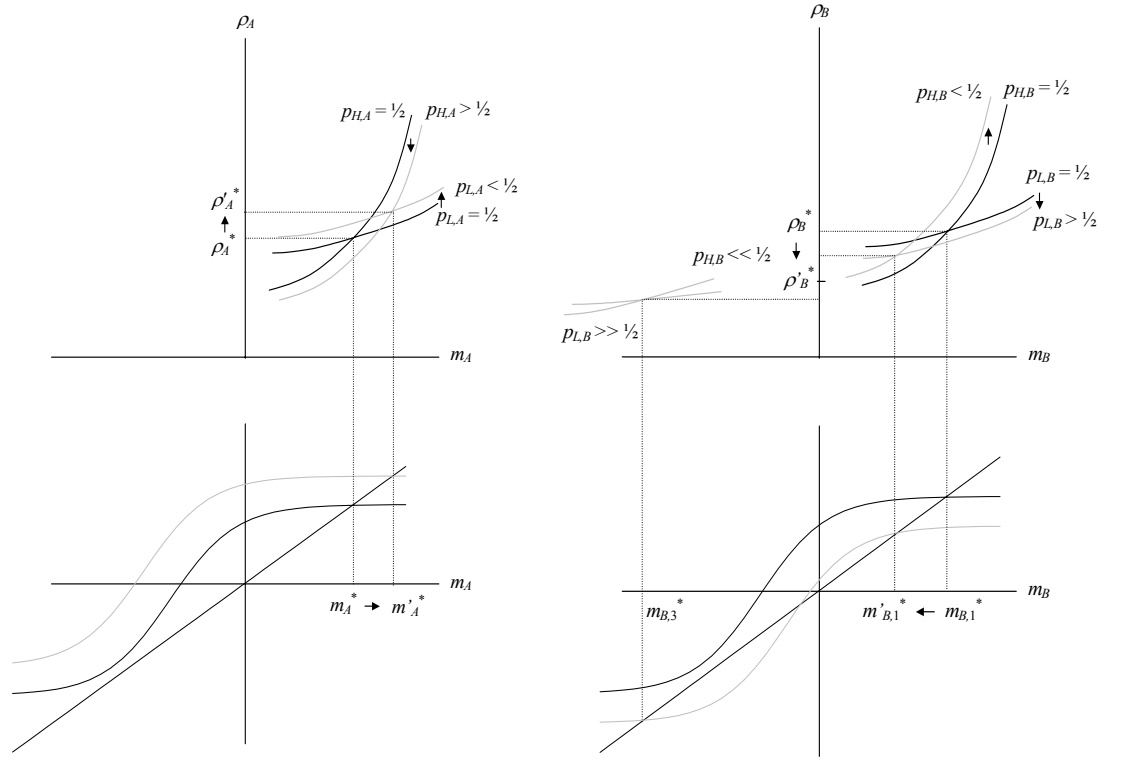


Figure 4. Adjustment in group A (left), and in group B (right).

²⁶Figure 4 assumes that the distribution of h is skewed enough, and h_L is small enough, to generate the critical \tilde{f}_{Hg} described in the corollary to proposition 3.

The new equilibrium with stratification is stable, since: (1) the single crossing property makes the process irreversible, and (2) given β_γ (which measures unobserved heterogeneity) the same equilibrium is necessarily reached from the other direction, i.e. if we start from a set of isoprobability curves generating a higher degree of stratification than the equilibrium one. The existence of a third, symmetric, equilibrium with stratification follows easily if we invert the roles of groups A and B in the thought experiment. Finally, it is clear that whatever the initial arrangement of the population across groups, i.e. whatever the initial membership probabilities - except $p_{hA} = p_{hB} = .5$ - the system will reach one of the two stable equilibria. This completes the proof of proposition 6, whose meaning is illustrated in figure 5

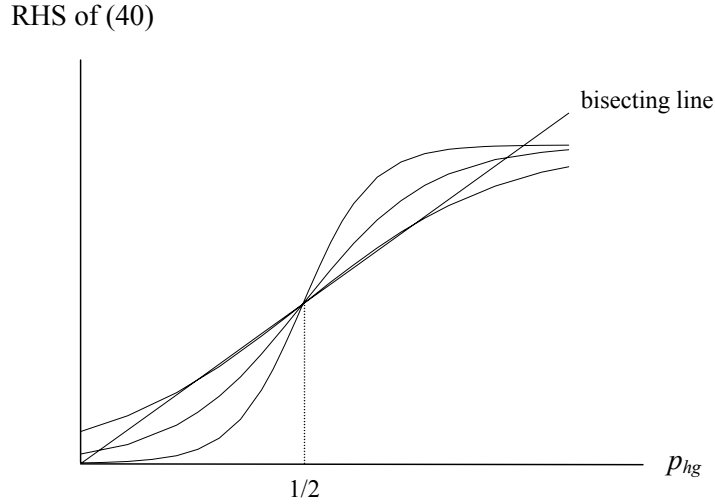


Figure 5. Fixed points of the membership probability function.

The different curves in figure 5 correspond to different preference parameters. Specifically, steeper curves, which generate higher degrees of equilibrium stratification, are associated with a larger value of the interactions parameter²⁷, J_g , and with smaller dispersion of the unobservable preferences over memberships, i.e. larger β_γ . The relation between equilibrium degree

²⁷Notice that when $J_g = 0$, the right hand side of (40) no longer depends on m_g , and so neither depends on p_{hg} . In this case the membership probability function becomes flat, and the only sorting equilibrium is integration.

of stratification and such parameters, can be illustrated formally, as follows. The total variation in membership probabilities when price and expected mean behavior change is:

$$dp_{hg} = \frac{\partial p_{hg}}{\partial \rho_g} d\rho_g + \frac{\partial p_{hg}}{\partial m_g} dm_g \quad (41)$$

where

$$\begin{aligned} \frac{\partial p_{hg}}{\partial \rho_g} &= -\beta_\gamma p_{hg} (1 - p_{hg}) \\ \frac{\partial p_{hg}}{\partial m_g} &= J_g \beta_\gamma p_{hg} (1 - p_{hg}) \tanh(h + J_g m_g) \end{aligned} \quad (42)$$

If we start at the integrated sorting equilibrium, $\varsigma = 0$, the equilibrium degree of stratification is equal to the variation in the degree itself. Defining $h^* = \arg \max_h \{|p_{h,A} - p_{h,B}|\}$, using definition (35) and the differential (41), we can compute the equilibrium degree of stratification as follows:

$$\begin{aligned} d\varsigma &= |dp_{h^*A} - dp_{h^*B}|_{p_{h^*A} = p_{h^*B}} \\ &= |\beta_\gamma [p_{h^*A} (1 - p_{h^*A}) (J_A \tanh(h^* + J_A m_A) dm_A + d\rho_A) \\ &\quad - p_{h^*B} (1 - p_{h^*B}) (J_B \tanh(h^* + J_B m_B) dm_B + d\rho_B)]| \end{aligned} \quad (43)$$

This expression suggests that the equilibrium degree of stratification is small if either of the following holds:

1. $J_g \rightarrow 0$ (i.e. interactions are weak), which means that when the endogenous social effect is negligible, group composition does not affect utility much, and so the incentive to segregate is weak.
2. $\beta_\gamma \rightarrow 0$ (i.e. ε_{ig} has a high dispersion), which means that changes in prices and expected mean behavior in a group do not affect membership probabilities significantly when individuals have a possibly strong unobservable preference for a certain group. In this case they might be unwilling to sell their membership, even for a very high price.

6 Welfare analysis

This section answers question (3): is the decentralized equilibrium efficient? To answer such question, we must figure out how a social planner would choose group compositions, i.e. f_{hg} , $h \in \Theta$, and $g \in G$. So, let's delegate the first-stage problem to a planner, who allocates the individuals across groups, and then lets individuals choose their behavior. Given the nature of the problem, and assuming the planner is utilitarian, the appropriate measure of social welfare, U , is the sum of equilibrium individual inclusive utilities, i.e. the aggregate value of memberships, before behavior is chosen, given that each individual will behave optimally:

$$\begin{aligned} U &= \sum_h \sum_g W_{hg}^* \\ &= \frac{1}{2} (f_{HA} W_{HA}^* + f_{LA} W_{LA}^* + f_{HB} W_{HB}^* + f_{LB} W_{LB}^*) \end{aligned} \quad (44)$$

The planner's problem is:

$$\max_{f_{HA}, f_{HB}} U \quad (45)$$

Such problem does not have a closed-form solution, since such a solution depends on $\partial m_g / \partial f_{HA}$, $g \in G$, and this partial itself has no closed form. However, a necessary condition for the pair $\{f_{HA}^{**}, f_{HB}^{**}\}$ to solve problem (45), is that it also solves:

$$\max_{f_{HA}, f_{HB}} \sum_z \mu(m_{g,z}^*) (m_{A,z}^*(f_{HA}) + m_{B,z}^*(f_{HA})) \quad (46)$$

where $z = 1, 2, 3$, to allow for multiple behavioral equilibria²⁸. In general, we cannot be sure that the decentralized sorting process leads to an efficient outcome, nor that it does not. However, when there exist sorting equilibria that generate multiple behavioral equilibria in one group, and the parameters are such that with zero stratification the behavioral equilibrium is unique, the

²⁸Given that $h > 0$, the condition is necessary for the following reason. We know that W_{hg} is increasing in $m_{g,z}^*$. But, while $m_{A,z}^*$ is increasing in f_{HA} , $m_{B,z}^*$ is decreasing in the same variable. Therefore, U cannot be at a maximum if the weighed sum of expected mean behaviors in the two neighborhoods is not at a maximum.

decentralized solution is definitely inefficient, since it generates a “social trap” which reduces the value of membership in a group. This can be summarized as follows:

Proposition 7 *An arrangement $\{f_{HA}, f_{HB}\}$ is inefficient if one group has multiple behavioral equilibria under this arrangement, and a unique behavioral equilibrium under any other arrangement. Therefore, if a critical concentration of H -types, $\tilde{f}_{Hg}(J_g)$, exists (see Corollary to Proposition 3), then the decentralized sorting process leads to an inefficient outcome.*

Proof. In case of multiple equilibria in one group g , $m_{g,z} < 0$ for $z > 1$, since $h > 0$. Therefore, U cannot be at a maximum if there exists another arrangement which generates unique behavioral equilibria in both groups. Q.E.D.

Corollary. If \tilde{f}_{Hg} exists (as defined in the Corollary to Proposition 3), the planner’s solution is in the interval $\left[\frac{1}{2}, \tilde{f}_{Hg}\right]$.

When the sorting process leads to an inefficient outcome, the planner can implement the optimal allocation through an appropriate system of membership taxes and subsidies.

7 A numerical exercise

The working of the model can be illustrated through a numerical exercise, which helps to overcome the problem of nonexistence of a closed form solution for the roots of equation (31). The equilibrium equations are numerically simulated, assuming $k_A = k_B$ and $J_A = J_B = J$.²⁹ The exercise considers the simplest case, i.e. $f_L = f_H$. The interesting case is that in which multiple behavioral equilibria are possible. Therefore, let’s pick the following values for the parameters: $H = 10$, $L = 2$, $J = 10$, where the parameters are all multiplied by $\beta = 1$, the parameter of the distribution of $\varepsilon_{ig\omega}$. These values

²⁹The first assumption means that the groups are identical in their exogenous characteristics: this allows isolation of the effect of their social composition. The second assumption means that interactions are global and of identical magnitude in each group.

are also such that at the integrated equilibrium both groups have a unique equilibrium $m_g^* \simeq 1$, but a critical value \tilde{f}_{Hg} exists, as defined in the corollary to proposition 3. Let's assume the following on the probability distribution $\mu(m_g^*)$ in case of multiple equilibria: individuals assign zero probability to the unstable equilibrium, and equal probabilities to the stable ones. Table 1 reports the results of such numerical exercise. The 1's in the first two rows are just approximations: in fact those numbers are below one and increase with the degree of segregation, as lemma 1 above states. It turns out that the value of $\tilde{f}(H)_g$ is close to 0.75. Below this value, the stable equilibrium (which depends on the value of β_γ) is very close to integration, although it cannot be fully integrated. Above it, the equilibrium can have a significant degree of segregation.³⁰ The stability

$m_{A,1}$	1	1	1	1	1	1	1
$m_{B,1}$	1	1	1	1	1	1	1
$m_{B,3}$	-	-.4960	-.5996	-.7007	-.8039	-.9156	-.9751
W_{HA}	20	=	=	=	=	=	=
W_{LA}	12	=	=	=	=	=	=
W_{HB}	20	12.52	12	11.5	10.99	10.51	10.36
W_{LB}	12	7.48	8	8.5	9.02	9.58	9.88
p_{HA}	.50	.75	.80	.85	.90	.95	.98
p_{LA}	.50	.25	.20	.15	.10	.05	.02
p_{HB}	.50	.25	.20	.15	.10	.05	.02
p_{LB}	.50	.75	.80	.85	.90	.95	.98
$\rho_A - \rho_B$	0	5.9993	5.9998	5.9994	5.9951	5.9576	5.8813
ζ	0	.5	.6	.7	.8	.9	.96
$\Delta U _{\beta_\gamma/\beta}$	-	-5.26	-4.8	-4.25	-3.58	-2.78	-2.27
β_γ/β	any	.742	.694	.693	.729	.833	1.036

Table 1. Results of the numerical exercise.

The exercise shows that substantial equilibrium stratification emerges as one group experiences a phase transition to a multiple equilibria regime. It also shows that a departure from the integrated equilibrium is never efficient

³⁰Of course the roles of the two neighborhoods can be inverted. In fact, every sorting equilibrium results in two symmetric sorting equilibria.

if it causes a phase transition. However, repeating the exercise with different parameter values shows that some degree of stratification can be efficient before a phase transition occurs. Stability of the equilibria, for given β_γ , can be verified noting that the expected mean behaviors in each column are generated by the memberships probabilities in the same column. So, different membership probabilities cannot be equilibrium probabilities and, given β_γ , they must converge to the values reported in each column, as iteration starting from different values of p_{hg} shows.

8 Econometrics

The unbiased identification of the endogenous social effect (i.e. of the parameter J) is a major concern in the econometric and empirical literature on social interactions (see Durlauf, 2003 sections 3 and 4, for a valuable overview). The detection of neighborhood effects, and the associated social multiplier, is important to answer delicate policy and organizational questions. For instance: does desegregation increase social welfare? Should classrooms be stratified along “intelligence”? Should workers be organized in skill-homogeneous teams? The purpose of this section is to show that the nested logit framework developed in this paper can be of help in overcoming problems of identification and selection-bias. One can directly estimate the nested logit model for the binary case. The procedure, based on two-stage maximum likelihood, is illustrated in detail by Ben-Akiva and Lerman (1985, pp. 295-299). This procedure identifies the parameters, but is subject to selection bias. Rather than discussing selection-correction in the logistic model, I will show how to exploit the theory to correct the “baseline”, linear-in-means, model. Such model describes the (continuous) behavior of an individual, ω_i , drawn from a group g :

$$\omega_i = a + cX_i + dY_g + Jm_g + \varepsilon_i \quad (47)$$

In this equation, X_i is a set of individual explanatory variables, Y_g a set of group specific exogenous characteristics (therefore d is the coefficient of the exogenous, or contextual, effect), and m_g is the *expected* behavior in group g . So, J is the coefficient of the endogenous social effect. This is the model studied in Manski’s (1993) seminal paper. The main econometric issues have been subsequently developed by Brock and Durlauf (2001b). As

for identification, notice that under rational expectations m_g can be rewritten as:

$$\begin{aligned} m_g &= E(\omega_i|Y_g) = a + cE(X_i|Y_g) + dY_g + Jm_g \\ &= \frac{a}{1-J} + \frac{c}{1-J}E(X_i|Y_g) + \frac{d}{1-J}Y_g \end{aligned} \quad (48)$$

Model (47) can be rewritten in reduced form:

$$\omega_i = \frac{a}{1-J} + cX_i + \frac{d}{1-J}Y_g + \frac{Jc}{1-J}E(X_i|Y_g) + \varepsilon_i \quad (49)$$

It is clear that the parameters in the behavioral equation (47) are not identified when the group mean of the individual characteristics, $E(X_i|Y_g)$, is a linear function of group characteristics, Y_g . In this case, the number of parameters that can be identified would be smaller than the number of parameters we need to estimate, i.e. (a, c, d, J) . In particular the main parameter of interest, J , is not identified. As for selection bias, notice that the estimation of a model like (47) requires orthogonality between the unobservables and the regressors, i.e.:

$$E(\varepsilon_i|X_i, Y_g, m_g) = 0 \quad (50)$$

But if individual i has chosen group g following some selection rule of the kind $R_i(X_i, Y_g, m_g, \varepsilon_{ig}) \geq 0$, then it is not reasonable to assume $E(\varepsilon_i|R_i \geq 0) = 0$. This is Heckman's (1979) selection problem: equation (47) describes behavior in the population, but the sample contains only those individuals who have chosen g , according to rule $R_i \geq 0$. This introduces selection-bias, a problem whose solution requires an estimate of $E(\varepsilon_i|R_i \geq 0)$, to be used as a selection-correction term. The theoretical model developed in the paper can be used to solve both problems, in three steps. First, consider inclusive utility of group g for individual i , in the continuous case:

$$W_{ig} = \log \int_{\omega_{\min}}^{\omega_{\max}} \exp(u_i(w, X_i) + Jm_g) dw \quad (51)$$

Second, consider the group selection rule implicit in equation (12):

$$g = \arg \max_{\nu} \bar{V}_{i\nu} \quad (52)$$

where $\bar{V}_{ig} \equiv \max_{\omega} V_i(g, \omega) = k_g - \rho_g + W_{ig} + \varepsilon_{ig}$ is expected maximum utility of group g . This rule, which rests on a maximization process, has the form we need to compute the selection-correction term. Assuming, for ease of exposition, that we are working with a sample partitioned into two groups, the correction term is:

$$E(\varepsilon_i | \bar{V}_{ig} - \bar{V}_{ig'} \geq 0) \quad g' \neq g$$

Finally, we use this term (which is a function of prices, exogenous group characteristics, and inclusive utilities) to correct for self-selection the linear-in-means model (47)³¹. The idea behind this procedure is that, as suggested by the theoretical model, inclusive utility affects preferences for groups when individuals consider the social effects at the second stage. The key to identification is the fact that inclusive utility is a nonlinear function of expected mean behavior³². Specifically, let's use the expected maximum utilities of the two groups to construct a latent variable Q_{ig} , such that group g is chosen if and only if $Q_{ig} > 0$:

$$Q_{ig} = q_1 (W_{ig} - W_{ig'}) + q_2 (k_g - k_{g'}) - q_3 (\rho_g - \rho_{g'}) + \eta_{ig} \quad (53)$$

$g' \neq g$, where η_{ig} , being the difference between two extreme value random variables, is logistically distributed. Assuming that η_{ig} and ε_i are jointly normal, we estimate $E(\varepsilon_i | Q_{ig} > 0)$ as follows (see Heckman, 1979, and Brock and Durlauf 2001b, 2003):

$$E(\varepsilon_i | Q_{ig} > 0) = \frac{\sigma_{\varepsilon\eta}}{\sigma_{\eta}} \frac{\phi(Q_{ig})}{\Phi(-Q_{ig})} \equiv H(m_g, k_g, \rho_g) \quad (54)$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are, respectively, the density and distribution functions of the standard normal. Therefore, the following is a model corrected for the

³¹Ioannides and Zabel (2002) were the first to use the nested multinomial logit model this way, although their procedure is not related to a general equilibrium framework.

³²See Brock and Durlauf (2001b, § 3.5-3.6) for a discussion of the relation between nonlinearities and identification.

fact that the sample self-selected into group g , provided that the alternative group g' was an available choice³³:

$$\omega_i = a + cX_i + dY_g + Jm_g + H(m_g, k_g, \rho_g) + \zeta_i \quad (55)$$

By construction, the residual error ζ_i is independent of all the other regressors³⁴. Notice that the estimation of this models requires nonlinear techniques, since the correction term is nonlinear in the interactions parameter, J . How can a model like (55) be useful? If the assumption of the paper is correct (i.e. if people sort taking into account social interactions), then the model can produce the unbiased identification of the interactions parameter, J . If the assumption is not correct, then the coefficient of the difference between inclusive utilities, q_1 , will not be statistically significant. In this case the nonlinearity that allows identification is lost, and identification breaks down. However, the theory I have proposed would be falsified.

9 Discussion and conclusions

I developed a framework for the theoretical and econometric analysis of choice subject to endogenous social effects, when individuals can sort themselves into groups, buying memberships. This is an investigation into an important and still underexplored area in social interactions economics. The model I have used exploits an idea due to Brock and Durlauf (2003), and exploited econometrically by Ioannides and Zabel (2002). The main findings concern: (1) the equilibrium relation between expected social interactions and expected group composition: the former induce equilibrium stratification; (2) the econometric implications of the theory, with respect to identification and the self-selection problems in the empirical detection of social interactions. The application of such a framework requires working out several weaknesses. First, the model relies heavily on parametric assumptions, concerning both the functional relations and the distributions of unobservables. Although

³³This is in fact a major complication, since one can observe choice but the not the choice set.

³⁴This way of solving identification and selection problems suffers from dependence on parametric assumptions, as stressed by Durlauf (2003) who advocates the importance of a non-parametric approach to selection correction.

such assumptions allow me to reach neat conclusions, they produce a substantial loss in terms of generality. A nonparametric development of the model is on my research agenda, notably in point of identification under weaker assumptions, following Brock and Durlauf (2004). Second, I have made extensive use of the rational expectations assumptions. Again, this simplifies the analysis a lot, and imposes a discipline on the model. The drawback of this shortcut is that it is very demanding in terms of substantive rationality, and this might raise issues of plausibility in applications to several interesting contexts, such as those in which the two stages of choice occur at distant points in time. Think of parents choosing a residential neighborhood, and their offspring making decisions about their human capital in presence of neighborhood effects. This example highlights a third weakness: I have assumed that preferences are aligned at the two stages of choice. In many applications, such as the parents-offspring one, preferences are not aligned if the decision-maker at the second stage is subject to social effects that the decision-maker at the first stage may dislike. For instance, parents may dislike the fact that their children conform to peers' behavior, while a formulation like (10) rules out such a possibility. The extension of the model to genuinely two-stage decision processes, with possible divergence in preferences, is an interesting research direction. Fourth, the theoretical model can account for all observed degrees of stratification, just varying the dispersion of the unobservables, which is not a very good property. However, I think what matters is the fact that the model predicts some degree of equilibrium stratification, and allows comparative statics. Fifth, the description of individuals has been quite naive. In particular, there is no budget constraint when buying memberships. Although, in a sense, the individual type works as a constraint, a detailed modeling of budget constraints would definitely improve the model. Finally, although the main advantage of the framework I have developed is that it leads to a corrected econometric model, there is an issue that requires more research. In the data we can observe chosen memberships, but not the choice set from which the observed individual membership was chosen. Uncertainty about the set of available groups undermines the procedure I have proposed, since the latter requires comparing inclusive utilities for all alternatives that were considered. Hints for a solution to this problem come from recent research on model uncertainty and model averaging (Brock, Durlauf and West, 2003), which allows to handle situations in which the true model, in this case the actual choice set, is unknown, although it belongs to a known set of models.

Even so, the model can be applied to study, theoretically and, in a tentative way, empirically, several interesting phenomena. I regard as particularly appealing the inquiry into the process of accumulation of human capital in presence of sorting and complementarities. A natural approach to this issue is the extension of models of household sorting and local spillovers in education (such as de Bartolome 1990, Bénabou 1996 and Durlauf 1996a,b). These models have assumed that children's human capital is affected by the community through some (not necessarily deterministic) function $g(x)$, where x is usually the proportion of H -type households. In this case children's choice, and the interactions among them, are not specified. The framework I have developed allows the modeling of kids' choice as well, and of the effect of expected (prospective) peers' decisions, rather than average human capital in the community only. This can be useful to highlight the role of sorting as a device to avoid allegedly "bad" interactions to one's own kids. Furthermore, the analysis has shown that extending the aforementioned models to endogenous interactions introduces the possibility of multiple behavioral equilibria, to be interpreted as a form of poverty trap: in presence of social effects, some communities end up encouraging inefficient behavior. I am confident that research in this direction, which offers the remarkable advantage of a directly implementable econometric model, can further enhance our ability to understand phenomena of social and economic inequality.

References

- [1] Akerlof, G. (1997), "Social Distance and Social Decisions", *Econometrica*, 65 (5), pp. 1005-1027.
- [2] Anderson, S., A. de Palma, and J.-F. Thisse (1992), *Discrete Choice Theory of Product Differentiation*, Cambridge Mass.: MIT Press.
- [3] Bayer, P., R. McMillan and K. Rueben (2003), "An Equilibrium Model of Sorting in an Urban Housing Market: the Causes and Consequences of Residential Stratification", mimeo, Yale University
- [4] Ben-Akiva, M. (1973), *Structure of Passenger Travel Demand Models*, Ph.D. dissertation, Dept. of Civil Engineering, MIT, Cambridge Mass.

- [5] Ben-Akiva, M. and S. Lerman (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, Cambridge Mass.: MIT Press.
- [6] Bénabou, R. (1993), "Workings of a City: Location, Education, and Production", *Quarterly Journal of Economics*, 108 (3), pp. 619-652.
- [7] Bénabou, R. (1996), "Equity and Efficiency in Human Capital Investment: the Local Connection", *Review of Economic Studies*, 63, pp. 237-264.
- [8] Borsch-Supan, A. (1990), "On the Compatibility of Nested Logit Models with Utility Maximization", *Journal of Econometrics*, 43, pp. 373-388.
- [9] Brock, W. and Durlauf, S. (2001a), "Discrete Choice with Social Interaction", *Review of Economic Studies*, 68, pp. 235-260.
- [10] Brock, W. and Durlauf, S. (2001b), "Interactions-Based Models", in *Handbook of Econometrics* 5, J. Heckman and E. Leamer eds., Elsevier, pp. 3297-3380.
- [11] Brock, W. and Durlauf, S. (2003), "Multinomial Choice with Social Interactions", in *The Economy as an Evolving Complex System III*, L. Blume and S. Durlauf, eds, Oxford University Press, forthcoming.
- [12] Brock, W. and S. Durlauf (2004), "Identification of Binary Choice Models with Social Interactions", *SSRI Working Paper 2004-2*, University of Wisconsin-Madison.
- [13] Brock, W., S. Durlauf, and K. West (2003), "Policy Evaluation in Uncertain Economic Environments", *Brookings Papers on Economic Activity*, 1/2003, pp. 235-322.
- [14] de Bartolome, C. (1990) "Equilibrium and Inefficiency in a Community Model with Peer Group Effects", *Journal of Political Economy*, vol. 98, 11, pp. 110-133.
- [15] Debreu, G. (1960), "Review of D. Luce, Individual Choice Behavior: A Theoretical Analysis", *American Economic Review*, 50, 186-188.
- [16] Durlauf, S. (1996a), "A Theory of Persistent Income Inequality", *Journal of Economic Growth*, 1, pp. 75-93.

- [17] Durlauf, S. (1996b), “Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality”, in Barnett, Gandolfo, and Hillinger eds., *Dynamic Disequilibrium Modeling*, Cambridge Mass.:Cambridge University Press, pp. 505-534.
- [18] Durlauf, S. (2003), “Neighborhood Effects”, SSRI working paper, University of Wisconsin-Madison (forthcoming in J.V. Henderson, and J.-F. Thisse eds., *Handbook of Regional and Urban Economics*, vol. 4).
- [19] Epple, D. (2003), “Modeling Population Stratification Across Locations: an Overview”, *International Regional Science Review*, 26, 2, pp. 189-196.
- [20] Epple, D., and G. Platt (1998), “Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Income”, *Journal of Urban Economics*, 43, pp. 23-51.
- [21] Evans, W., W. Oates and R. Schwab (1992), “Measuring Peer Group Effects: A Study of Teenage Behavior”, *Journal of Political Economy*, vol. 100, 51, pp. 966-991.
- [22] Glaeser, E., and J. Scheinkman (2001), “Measuring Social Interactions”, in *Social Dynamics*, S. Durlauf and H. P. Young, eds., Cambridge Mass.: MIT Press, pp. 83-131.
- [23] Glaeser, E., B. Sacerdote, and J. Scheinkman (1996), “Crime and Social Interactions”, *Quarterly Journal of Economics*, 111, 2, pp. 507-548.
- [24] Heckman, J. (1979), “Sample Selection Bias as a Specification Error”, *Econometrica*, 47(1), 153-161.
- [25] Herriges, J., and C. Kling (1996), “Testing the Consistency of Nested Logit Models with Utility Maximization”, *Economic Letters*, 50, pp. 33-39.
- [26] Ichino, A., and A. Falk (2003), “Clean Evidence on Peer Pressure”, *mimeo*, EUI.
- [27] Ioannides, Y., and J. Zabel (2002), “Interactions, Neighborhood Selection, and Housing Demand”, *mimeo*, Tufts University.

- [28] Manski, C. (1993), “Identifications of Endogenous Social Effects: The Reflection Problem”, *Review of Economic Studies*, 60, 531-542.
- [29] Manski, C. (1995), *Identification Problems in the Social Sciences*, Cambridge Mass.: Cambridge University Press.
- [30] McFadden, D. (1978), “Modeling the Choice of Residential Location”, in *Spatial Interaction Theory and Planning Models*, A. Karlqvist, L. Lundqvist, F. Snickars and J. Weibull eds., North-Holland, pp. 75-96.
- [31] Milgrom, P. and J. Roberts (1990), “Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities”, *Econometrica*, 58, 6, pp. 1255-1277.
- [32] Moffitt, R. (2001), “Policy Interventions, Low-Level Equilibria, and Social Interactions”, in *Social Dynamics*, S. Durlauf and H. P. Young, eds., Cambridge Mass.: MIT Press, pp. 45-82.
- [33] Sacerdote, B. (2001), “Peer Effects with Random Assignments: Results from Dartmouth Roommates”, *Quarterly Journal of Economics*, 116, 2, pp. 681-704.
- [34] Samuelson, L. (2004), “Information-Based Relative Consumption Effects”, *Econometrica*, 72, 1, pp. 93-118.
- [35] Train, K. (2003), *Discrete Choice Methods with Simulation*, Cambridge Mass.: MIT Press.
- [36] Zimmerman, D. (2003), “Peer Effects in Higher Education: Evidence from a Natural Experiment” *Review of Economics and Statistics*, 85, 1, pp. 9-23.

10 Technical appendix

Derivation of inclusive utility and group choice probability.

The argument here follows Ben-Akiva and Lerman (1985, pp. 287-288). An alternative source is Anderson et al. (1992, pp. 60-61).

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 x_1 and x_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

$$\log [\exp(x_1) + \exp(x_2)]$$

Therefore, $\max_{\omega} (h_i \omega_i + J_g \omega_i m_{ig}^e + \varepsilon_{ig\omega})$ is EV distributed, with position parameter equal to

$$\log \sum_{\omega} \exp (h_i \omega_i + J_g \omega_i m_{ig}^e)$$

and scale parameter equal to 1.

Since the expected value of an EV random variable is its mode (the position parameter) plus Euler constant, γ , divided by the scale parameter, we have:

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

The probability that individual i chooses group g is:

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

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

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

So, p_{ig} can be rewritten as follows:

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

At the end of Section 2 I have assumed that ε_{ig} is distributed such that the maximum of $V_i(g, \omega)$ given a group, i.e. $\varepsilon_{ig} + \tilde{\varepsilon}_{ig}$ is EV distributed with parameter β_γ . Therefore, the same argument leading to equation (16) for behavior probabilities apply here and leads to group probabilities (18).

Proof of Proposition 1.

Since $\varepsilon_{i,g}$ is Extreme Value, it is continuously distributed. Therefore, $p_{i,g}$ is continuous and differentiable in $(\rho_g, \rho_{g'}) \equiv \boldsymbol{\rho}$, $g \neq g'$.

Since $I_g^d(\boldsymbol{\rho}) = \sum_i p_{i,g}$, $I_g^d(\boldsymbol{\rho})$ is continuous and differentiable too. Equilibrium requires $I_g^d(\boldsymbol{\rho}) = I/2$, therefore we can define $z_g(\boldsymbol{\rho}) \equiv I_g^d - I/2$ as the excess demand for membership in group g . $z_g(\boldsymbol{\rho})$ is necessarily continuous, and:

$$\begin{aligned}
\lim_{\rho_g \rightarrow +\infty} z_g(\boldsymbol{\rho}) &= -\frac{I}{2} < 0 \\
\lim_{\rho_g \rightarrow -\infty} z_g(\boldsymbol{\rho}) &= \frac{I}{2} > 0
\end{aligned}$$

By continuity of $z_g(\boldsymbol{\rho})$, and given its limit behavior, there must exist a ρ_g^* such that $z_g(\rho_g^*, \rho_{g'}) = 0$, for any $\rho_{g'}$. This holds for $g \in G$. Therefore there must exist a pair (ρ_A^*, ρ_B^*) such that $z_A(\boldsymbol{\rho}) = z_B(\boldsymbol{\rho}) = 0$, i.e. a set of prices clearing the memberships market. This proves existence.

To prove uniqueness, it is sufficient to show that $z_g(\boldsymbol{\rho})$ is monotonic in ρ_g , i.e. that it crosses the horizontal axis only once. This is in fact the case:

$$\frac{\partial z_g(\boldsymbol{\rho})}{\partial \rho_g} = \frac{\partial \sum_i p_{i,g}}{\partial \rho_g} = \sum_i p_{i,g} (p_{i,g} - 1) < 0$$

Q.E.D.

Proof of Proposition 3 and its Corollary

The proof of both Proposition 3 and its Corollary follow from the properties of the $\tanh(\cdot)$ function. Since there's no closed-form solution for its fixed point, the proof cannot be given algebraically.

Concerning the Proposition: by continuity, it is clear from figure 2 that there exists a critical value \tilde{h}_g for the linear combination of the two possible types, H and L , beyond which the function crosses the 45 degrees line three times. The threshold depends inversely on both f_{Hg} and J_g , because given that $H > L$, a higher f_{Hg} means a higher function at each point of its range, and the higher J_g the steeper the function around the vertical axis.

Concerning the Corollary: when $H > \tilde{h}_g > L$, by definition of \tilde{h}_g , a group made of H -type individuals only ($f_{Hg} = 1$) has a unique equilibrium. And a group made of L -type individuals only ($f_{Hg} = 0$) has multiple equilibria. Therefore, the continuity of the $\tanh(\cdot)$ function guarantees that the threshold \tilde{f}_{Hg} exists, and that is increasing in the magnitude of interactions. Q.E.D.

Proof of Lemma 1

The proof follows the argument given above to prove proposition 3. Equation (30) states that, in equilibrium:

$$m_g = \sum_h f_{hg} \tanh(h + J_g m_{hg})$$

Let's impose the approximation $m_g \simeq m_{Hg} \simeq m_{Lg}$ and rewrite the equation as:

$$m_g = f_{Hg} \tanh(H + J_g m_g) + (1 - f_{Hg}) \tanh(L + J_g m_g)$$

Since $H > L$, it follows from the properties of the $\tanh(\cdot)$ function that a larger f_{Hg} means a higher function at each point of its range. Therefore, any stable equilibrium will be located at a higher level.

Q.E.D.