Neighborhood Effects: Accomplishments and Looking Beyond them[1]

Yannis M. Ioannides* **and Giorgio Topa****

Presented at the

"Journal of Regional Science 50th Anniversary Conference"

Federal Reserve Bank of New York

New York, April 23–24, 2009

`This Draft: June 28, 2009`

ABSTRACT

The paper addresses the empirical significance of the social context in economic decisions. Decisions of individuals who share spatial and social milieus are likely to be interdependent, and econometric identification of social effects poses intricate data and methodological problems, including dealing with self-selection in spatial and social groups. It uses a simple empirical framework to introduce social interactions effects at different levels of aggregation, and examines estimation problems into linear models, the impact of self-selection and of non-linearities. It also examines neighborhood effects in job matching and proposes a research agenda that offers new techniques and data sources.

JEL codes: C00, C81, J6, R22, Z13.

Keywords: Neighborhood effects, social interactions, social networks, social effects, self-selection, neighborhood choice, social multiplier, spatial effects.

---

[1]We thank Gilles Duranton and the participants at the JRS 50th Anniversary Conference for very helpful comments and stimulating discussion.

* Department of Economics, Tufts University, Medford, MA 02155. +1 617 627 3294 `yannis.ioannides@tufts.edu`

** Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045. +1 212 720 5497 `giorgio.topa@ny.frb.org`

Contents

1 Introduction

The empirical economics literature on social interactions addresses the significance of the social context in economic decisions. Decisions of individuals who share spatial and social milieus are likely to be interdependent. Recognizing the nature of such interdependence in a variety of conventional and unconventional settings and measuring empirically the role of neighborhood effects, or more generally, of social interactions, poses complex econometric questions. Their resolution may be critical for a multitude of phenomena in economic and social life and for matters of public policy.

Broadly speaking, social interactions arise when individuals (or households) affect each other's decisions, preferences, information sets, and outcomes, directly rather than indirectly through markets. These interpersonal effects are known as *endogenous* social effects when own *decisions* and those of others in the same social milieu are inter-dependent. For instance, this occurs when individuals care not only about their own outcomes but also about the outcomes of others, such as the kinds of cars owned by their friends or their education decisions. Individuals may also care about personal characteristics of others, that is whether they are young or old, black or white, rich or poor, trendy or conventional, and so on, and about other attributes of the social milieu that may not be properly characterized as deliberate decisions of others. Such effects are known as exogenous social or *contextual* effects. We address below the particular difficulties that they pose for estimation. In addition, individuals in the same or similar social settings may act similarly because they share common unobservable factors or face similar institutional environments. Such an interaction pattern is known as *correlated* effects. This terminology is due to Manski (1993), who emphasized the difficulty of separately identifying endogenous from contextual effects in linear-in-means models, as well as identifying social effects — whether endogenous or exogenous (contextual) — from correlated effects.

Social interdependencies emerge naturally if individuals share a common resource or social space in a way that is not paid for but still generates constraints on individual action. Theorizing in this area lies at the interface of economics, sociology and psychology and is often imprecise. Its spatial aspects put it firmly in the realm of regional science. Terms like social interactions, neighborhood effects, social capital and peer effects are often used as synonyms although they may have different connotations. Empirical distinctions between endogenous, contextual and correlated effects are critical for policy analysis because of the presence of a "social multiplier" as we see in more detail further below.

Joint dependence among individuals' decisions *and* characteristics within a spatial or social milieu is complicated further by the fact that in many circumstances individuals in effect *choose* their own social context. That is, individuals in choosing their friends and/or their neighborhoods also choose their neighborhood effects as well. Such choices involve information that is in part unobservable to the analyst, and therefore require making inferences among the possible factors which contribute to decisions [Brock and Durlauf (2001) and Moffitt (2001)].

The remainder of this paper merely touches on what is truly a vast and continuously growing literature in the social sciences. It is organized as follows. Section 2 below uses a simple empirical framework to introduce social interactions effects and Section 3 the estimation problems they pose in a variety of settings. It examines in a bit more detail, and for the purpose of an example, a particular estimation approach that rests on the notion of the social multiplier. This lends itself naturally to use of data at different levels of aggregation. Section 4 emphasizes how nonlinearities improve the prospects for identification and demonstrates their significance in the case of self selection intro groups. Section 5 takes up neighborhood effects in job matching. This is a particularly interesting area because local labor markets may involve both spatial and social interdependence. Section 6 concludes with a brief agenda about research that we think is quite likely to bear fruit.

2 A Simple Empirical Framework

We consider a static setup, for simplicity. Individual i's action y_i is a linear function of a vector of observable individual characteristics, $\mathbf{x}_i$, of a vector of contextual effects, $\mathbf{z}_{\nu(i)}$, which describe i's neighborhood $\nu(i)$, and of the expected action $\frac{1}{|\nu(i)|}\sum_{j\in\nu(i)}\mathcal{E}[y_j]$ among the members of i's neighborhood $\nu(i)$. That is:

$$y_i = \alpha_0 + \mathbf{x}_i\alpha + \mathcal{E}[\mathbf{z}_{\nu(i)}|\Psi_i]\theta + \beta\frac{1}{|\nu(i)|}\sum_{j\in\nu(i)}\mathcal{E}[y_j|\Psi_i] + \epsilon_i, \tag{1}$$

where Ψ_i are the attributes of i's neighborhood $\nu(i)$.

The endogenous social effect is defined with respect to the expectation of a contemporaneous endogenous variable. Abstracting at the moment from the issue that individual i may have deliberately chosen her neighborhood, $\nu(i)$, (which we take up in section 4.1 below) and stating that conditional on individual characteristics, contextual effects and the event that i is a member of neighborhood $\nu(i)$ the expectation of ϵ_i is zero, allows us to focus on the estimation of such models. Assume that we are in a *social equilibrium* and that individuals hold *rational expectations* over $\mathcal{E}[y_j|\Psi_i]$. That is, individuals' expectations are confirmed; they are equal to what the model predicts. So, taking the expectations of both sides of (1) and setting the expectation of y_i equal to $\frac{1}{|\nu(i)|}\sum_{j\in\nu(i)}\mathcal{E}[y_j|\Psi_i]$ allows to solve for this expectation. Substituting back into (1) yields a *reduced form*, an expression for individual i's outcome in terms of all observables $(\mathbf{x}_i, \mathbf{x}_{\nu(i)}, \mathbf{z}_{\nu(i)})$, where $\mathbf{x}_{\nu(i)}$ denotes the neighborhood average of personal attributes $\mathbf{x}_i$:

$$y_i = \frac{\alpha_0}{1-\beta} + \mathbf{x}_i\alpha + \frac{\beta}{1-\beta}\alpha\mathbf{x}_{\nu(i)} + \frac{\theta}{1-\beta}\mathbf{z}_{\nu(i)} + \epsilon_i. \tag{2}$$

We explore the intuition behind (1) by taking y_i to be an educational attainment. Own socioeconomic characteristics, $\mathbf{x}_i$, typically do affect educational attainment. The socioeconomic characteristics of adult neighbors, including measures of economic success, are often used as contextual effects and are included in $\mathbf{z}_{\nu(i)}$. They may represent *role model effects.* In contrast, the effect of educational attainment by one's peers in schools and neighborhoods, an endogenous social effect, is an example of a *peer group effect.* Note those effects are associated with distinct populations, which would be highlighted in a more detailed model.

A comparison of the model of (1) and its reduced form (2) allows one to demonstrate that endogenous social effects generate feedbacks which magnify the effects of neighborhood characteristics. That is, the effect of $\mathbf{z}_{\nu(i)}$ is $\frac{\theta}{1-\beta}$, and thus magnified, if $0 < \beta < 1$, relative to θ. It also confirms why it is tempting for empirical researchers to study individual outcomes as functions of all observables.

Consider the effect on the academic performance of a particular medical student of the presence of women in the classroom, measured in per cent. This problem is addressed by Arcidiacono and Nicholson (2005). According to (1), this effect is given by θ. However, this would ignore the fact that there is such an effect on every other student, conditional on their characteristics. Therefore, the effect is magnified, exactly as suggested by Equ. (2), and is now given by $\theta + \beta\theta + \beta\theta^2 + \ldots = \frac{\theta}{1-\beta}$.

Following the pioneering work of Datcher (1982), a great variety of individual outcomes have been studied in the context of different notions of neighborhoods. Deriving causal results requires great care with econometrics in addition to suitable data. The model of Equ. (1) is the bare minimum of interactions that we need in order to express essential complexities of social interactions. In practice, empirical researchers have dealt with models that are considerably more complex than (1). E.g., it is possible that the marginal effect of a neighbor's actions may depend on neighborhood characteristics. This can be expressed by a term $\mathbf{z}_{n(i)} \frac{1}{|\nu(i)|} \sum_{j\in\nu(i)} \mathcal{E}[y_j|\Psi_i]$ in (1). See Section 4 below. Linearity obscures the richness that comes with nonlinear social interactions models, like multiplicity of equilibria [Brock and Durlauf (2001)]; explicitly allowing for multiple equilibria in the estimation of social interactions effects yields striking and at times counter-intuitive policy implications [Bisin *et al.* (2009)].

3 Econometric Identification

Including as contextual effects only neighborhood averages of individual effects, $\mathbf{z}_{\nu(i)} \equiv \mathbf{x}_{\nu(i)}$, may cause failure of identification of endogenous separately from exogenous interactions. That is, we may not be able to estimate separately coefficients β and θ by means of a linear model like (1). Manski (1993) terms this the *reflection* problem: it arises because the direct effect of the social context variables $\mathbf{z}_{\nu(i)}$ shows up together with the indirect effect, as reflected through the endogenous effect represented by $\frac{1}{|\nu(i)|} \sum_{j\in\nu(i)} \mathcal{E}[y_j|\Psi_i]$. To see this, consider regressing individual outcomes on neighborhood averages of individual characteristics as contextual effects. Equ. (2) suggests that we may estimate the combined effect $\frac{\alpha\beta+\theta}{1-\beta}$. A statistically significant estimate of the coefficient of $\mathbf{x}_{\nu(i)}$ in a reduced form regression like (2) allows a researcher to infer that at least one type of social interaction is present: either β, or θ, or both are nonzero. Therefore, partial identification is possible for some type of social effect.

If the underlying economic model makes it possible to exclude some of neighborhood averages of individual covariates, then the model parameters may be identified. More precisely, for the identification of (1), the vector $\mathbf{x}_{\nu(i)}$ must be linearly independent of $(1, \mathbf{x}_i, \mathbf{z}_{\nu(i)})$. It is thus necessary that there be at least one element of $\mathbf{x}_{\nu(i)}$ whose group level average is not a causal effect and therefore not included in $\mathbf{z}_{\nu(i)}$. This rather high bar for identification is a direct consequence of linearity of the endogenous social effect in the behavioral model. The importance of precision for identification and the usefulness of economic models is demonstrated by the work of Cohen-Cole and Fletcher (2008), who take on sociologists on the identification of endogenous versus contextual effects in the context of obesity.

3.1 The Social Multiplier

The fact that endogenous social interactions help amplify differences in the average neighborhood behavior across neighborhoods can itself serve as basis for identification. Glaeser *et al.* (2003) use patterns in the data to estimate a *social multiplier*. For a change in a particular fundamental determinant of an outcome, this is defined as the ratio of a total effect, which includes a direct effect to an individual outcome plus the sum total of the indirect effects through the feedback from the effects on others in the social group, to the direct effect. An estimate of the multiplier can be seen as the ratio of the coefficient associated to $\mathbf{x}_{\nu(i)}$ in a "group level" regression of neighborhood outcomes on neighborhood attributes ($y_{\nu(i)}$ on $\mathbf{x}_{\nu(i)}$), to the "individual level" coefficient associated to $\mathbf{x}_i$ when regressing y_i on $\mathbf{x}_i$. Glaeser *et al.* (2003) show that this ratio tends to $1/\left[(1-\beta)\left(1+\sigma\beta\right)\right]$ as the number of observations grows large, where $\sigma = Var(\mathbf{x}_{\nu(i)})/Var(\mathbf{x}_i)$ is the portion of the variation in individual attributes that is due to the group level variation. Therefore, one can obtain an estimate of β, the endogenous social effect parameter of interest, from the ratio of group level to individual level regression coefficients and an estimate of σ.

This approach must deal, in practice, with dependence across decisions of individuals belonging to the same group, which is implied by non-random sorting in terms of unobservables. Specifically, if educated people prefer to have other educated people as neighbors, the effect of one person's education (in an individual-level regression) will overstate the true impact of education because it includes spillovers. So, with sorting on observables and positive social interactions, the individual level coefficient will overstate the true individual level relationship and the estimated social multiplier will tend to underestimate the true level of social interactions. On the other hand, correlation between aggregate observables and aggregate unobservables will cause the measured social multiplier to overstate the true level of social interactions.

This approach is particularly useful in delivering a range of estimates for the endogenous social effect and when individual data are hard to obtain, as in the case of crime data. For example, Glaeser, Sacerdote and Scheinkman (1996) motivate their study of crime and social interactions by the extraordinary variation in the incidence of crime across US metropolitan areas over and above differences in fundamentals. If social interactions in criminal behavior are present, variations in observed outcomes are larger than what would be expected from variations in underlying fundamentals. This is because positive social interactions generate positive spatial correlations across individuals in a given metropolitan area, which in turn raise the overall variance of crime across metro areas.

Glaeser, Sacerdote and Scheinkman (2003) report results using a multiplier-based approach with three different outcomes. One is fraternity/sorority participation by students in Dartmouth College. This setting has the advantage that students are randomly assigned to residences in Dartmouth College, in other words there is no sorting. So, aggregating at the room, floor and dormitory level allows these researchers to apply the multiplier technique, thanks to random group assignments. The impact of having drunk beer in high school on the probability of fraternity/sorority participation rises with the level of aggregation, exactly as the model predicts. This allows the authors to estimate the magnitude of the endogenous social interaction effect.

In practice, the authors estimate the social multiplier by first estimating individual coefficients associated with specific attributes, and then forming a predicted aggregate outcome based on aggregate characteristics and those individual-level coefficients. They then regress actual aggregate outcomes on these predicted aggregate outcomes; the coefficient from this final regression tends to $1/\left[(1-\beta)\left(1+\sigma\beta\right)\right]$ as N grows large and can thus be used to recover an estimate of the social interaction parameter β. This technique is also applied to crime data, as well as data on wages and human capital variables.

Graham (2008) also exploits the idea of excess variation across groups (e.g., classrooms) to identify social interaction effects. The logic is quite intuitive. The within-group sample variance of outcomes provides an estimate for the variance of individual heterogeneity (e.g., student ability across classrooms). The latter estimate, together with the between-group sample variance, can be used to estimate the variance of any group-level heterogeneity (excess variance). This excess variance can have two sources: the standard one due to group-level heterogeneity (e.g., teacher quality) and that coming from variation in peer quality across groups. If data on two or more subpopulations of groups are available, then a valid test for social interactions can be performed. In fact, if the distribution of group-level heterogeneity is the same across subpopulations, whereas that of peer quality differs, then it is possible to separately identify social interaction effects within classrooms.

4 Nonlinearities

The linear-in-means model described above is a very special case, however rigorously defined it might be. If dependence of outcomes on their determinants is inherently nonlinear, then full identification may be possible. This seems to be the case in the marriage market, an important urban market. Let y_i in Equ. (1) be individual i's propensity to marry in a given time period.

Individuals are affected by the marriage rate in their residential communities either because they are conformists, or because of the mechanics of marital matching. That is, higher marriage rate leaves fewer unmatched people, which affects one's own matching prospects negatively — fewer qualified people being left. The reduced flow of potential partners makes one less choosy. The net of these two effects is the endogenous social effect with coefficient β. On the other hand, someone might be more apt to marry if s/he lives in a community of wealthier potential matches, because such matches are more likely to marry and make more attractive spouses. This is the contextual effect, with coefficient θ. Further, people in communities, as defined by Ψ_i, that hold marriage in high regard are more likely to marry, but this is not known to the econometrician. That is, $\mathcal{E}[\epsilon_i|\Psi_i] = \Psi_i D \neq 0$: the correlated effect is nonzero and would appear as bias in the error of (1), where D is a vector of parameters, as we will clarify shortly.

Drewianka (2003) explicitly considers the two sides of the marriage market in each residential community as separate but interrelated "communities" where membership is exogenous. The marriage rate is

$$2\frac{\#|\text{matches}|}{\#|\text{male}| + \#|\text{female}|} = \left(\frac{\#|\text{matches}|}{\#|\text{group}_j|}\right)\left(\frac{\#|\text{group}_j|}{\#|\text{male}| + \#|\text{female}|}\right) = y_j r_j, \ j \in \{\text{male}, \text{female}\}$$

where r_j denotes the fraction of the marriage market attributed to group j. Taking expectations of both sides of (1) by considering it separately for each group j, using the above expression for the group-specific marriage rate, which is the endogenous social effect here, and solving for the endogenous social effect yields:

$$\mathcal{E}[y_j|\mathbf{z}_j, \Psi_j] = \mathcal{E}\left[\mathbf{x}_j\right]\alpha + \frac{1}{1 - 2\beta_j r_j}\left(\alpha_0 + \mathcal{E}\left[\mathbf{z}_j\right]\theta + \Psi_j D\right), \tag{3}$$

where we have simplified notation to denote by Ψ_j the vector of characteristics of individual i's group, by $\mathcal{E}\left[\mathbf{x}_j\right]$ the mean characteristics of group j, and by $\mathcal{E}\left[\mathbf{z}_j\right]$ the mean contextual effect. When we substitute back into Equ. (1), the resulting model is identified, because r_j varies across markets.

So, what is different here from Manski's reflection problem is that the endogenous social effect in an individual man's propensity to marry is not the marriage rate among men, the mean of the individual outcome. Instead, the endogenous social effect is a function of the *aggregate* marriage rate in the community. So, there is an additional source of variation: the greater the number of male potential marriage partners, r_{male} above, the lower the probability that a match will occur. The opportunity cost of getting married is larger when it is easier to search for potential alternative partners. In communities where most people are single, marriage rates are lower — not just algebraically, but also causally. Technically, identification of the endogenous social effect (the impact of marriage rates in one's group on an individual's likelihood of marrying) is possible because the relative sizes of the two sides of the market vary across communities and genders, adding one extra restriction to the estimation model, thus allowing resolution of the identification problem.

4.1 Identification of Social Interactions with Self Selection to Groups and Sorting

Roommates seem to be randomly allocated in Dartmouth College, a fact that allows Sacerdote (2001) to offer one of the most interesting cases of identification of social interactions effects. However, in general, presence of non-random sorting in terms of unobservables is a major challenge for the econometric identification of social interactions models. Choosing the group one belongs to implies that the random shock in the RHS of (1) may not be independent of the other regressors. That is, deliberate choice of neighborhood $\nu(i)$ by individual i suggests that the unobserved elements in the actions of individuals who have chosen the same neighborhood (or social group, more generally) are not independent of one another. Conditional on their characteristics, different individuals might still be influenced by unobservable common factors, rendering $\mathcal{E}[\epsilon_i|\mathbf{x}_i, \mathbf{z}_{\nu(i)}; \Psi_i; i \in \nu(i)] \neq 0$.

We formalize this notion by supposing that the evaluation of the attractiveness of a neighborhood ν may be expressed in terms of an unobservable "latent" quality variable $Q^*_{i,\nu}$.[2] That is, individual i evaluates neighborhood ν by means of observable attributes $W_{i,\nu}$ which enter with weights ζ, and unobservable component $\vartheta_{i,\nu}$:

$$Q^*_{i,\nu} = \zeta W_{i,\nu} + \vartheta_{i,\nu}. \tag{4}$$

Random shocks ϵ_i and $\vartheta_{i,\nu}$ are assumed to have zero means, conditional on (are orthogonal to) regressors $(\mathbf{x}_i, \mathbf{z}_{\nu(i)}, W_{i,\nu})$, across the population. If individual i were to choose

[2]The specification of the neighborhood quality index need not be arbitrary. It could be based on an underlying utility index, from which the quantity decision y_i also emanates [Ioannides and Zabel (2008)].

the neighborhood which affords her the highest possible evaluation, then the respective shocks would no longer have zero means. Once parametric assumptions are made about the joint distribution of $(\epsilon_i, \vartheta_{i,\nu})$, conditional on choosing neighborhood $\nu(i)$, an expression for $\mathcal{E}[\epsilon_i|\mathbf{x}_i, \mathbf{z}_{\nu(i)}; \Psi_i; i \in \nu(i)]$ may be obtained as proportional to a function $\delta(\zeta W_{i,\nu(i)})$, so that (1) may be rewritten as:

$$y_i = \alpha_0 + \mathbf{x}_i\alpha + \mathbf{z}_{\nu(i)}\theta + \beta\frac{1}{|\nu(i)|}\sum_{j\in\nu(i)}\mathcal{E}[y_j|\Psi_i] + \kappa\delta(\zeta W_{i,\nu(i)}) + \xi_i. \tag{5}$$

Combining information on the discrete choice of neighborhood problem (4) with information on the continuous outcome decision allows us to estimate such models.[3] The additional regressor $\delta(\hat{\zeta} W_{i,\nu(i)})$ in (5), where $\hat{\zeta}$ is obtained from the estimation of (4), even if it were to also include $\mathbf{z}_{\nu(i)}$, is generally nonlinear and therefore linearly independent of $(1, \mathbf{x}_i, \mathbf{z}_{\nu(i)})$. So in sum, if it is possible to estimate the neighborhood selection rule, then correction for selection bias via the mean estimated bias (the so-called Heckman correction term) introduces an additional regressor $\delta(\hat{\zeta} W_{i,\nu(i)})$ in the RHS of (1) whose neighborhood average is not in the RHS of Equ. (1) - in other words, it is not a causal effect.

Econometrically speaking, this approach supplies instruments that enable the identification of the model: in the reduced form regression (2) two new terms are introduced, $\kappa\delta(\zeta W_{i,\nu(i)})$ and $\kappa\beta/(1-\beta) \cdot \mathcal{E}[\delta(\zeta W_{i,\nu(i)})|i \in \nu(i)]$, but only one new parameter, κ. This allows for identification as long as $\delta(\zeta W_{i,\nu(i)})$ and $\mathcal{E}[\delta(\zeta W_{i,\nu(i)})|i \in \nu(i)]$ are not linearly dependent; this condition will be satisfied as long as there is within neighborhhood variation in $\delta(\zeta W_{i,\nu(i)})$. Brock and Durlauf (2001), who were the first to make this point, also extend it to duration data, and Sirakaya (2006) applies it to a study of recidivism.

Empirically, researchers have looked for policy experiments involving re-assignments of individuals to groups (schools or neighborhoods) as a way to address the problem of endogenous neighborhood choice. For instance, Boston's Metropolitan Council for Educational Opportunities (METCO) program is a long-standing voluntary desegregation program. The

[3]Brock and Durlauf (2001) and Durlauf (2004) provide more details on the econometric properties of the estimation process. Ioannides and Zabel, *op. cit.*, is an application of these methods on neighborhood selection and housing demand using confidential data from the US Census and the American Housing Survey.

program assists mainly black inner-city kids from Boston public schools in commuting to and enrolling in mainly white (and more prosperous) suburban Boston communities that accommodate them in their public schools. Angrist and Lang (2004) show, in seeking to evaluate the program, that the receiving school districts, which have higher mean academic performance than the sending ones, do experience a mean decrease due to the program. However, they also show that the effects are merely "compositional", and there is little evidence of statistically significant effects of METCO students on their non-METCO classmates. Analysis with micro data from one receiving district (Brookline, Massachusetts) generally confirms this finding, but also produces some evidence of negative effects on *minority* students in the receiving district.

METCO is noteworthy as a social experiment; it was initiated in 1966 by civil rights activists seeking to bring about de facto desegregation of schools. Lack of evidence of negative peer effects is particularly useful for informing desegregation policy. Still, it is a voluntary program for both sides, making self-selection a problem. There is self-selection at the individual level, by parents and children, and at the receiving school district level, by the political process that funds the program. There are numerous factors germane to selection, which are specific to how welcome the program was by each receiving school district, which is administered academically and fiscally by its respective community.

To overcome the selection issues, several papers have studied neighborhood effects by looking at the consequences of randomized assignment of residents of low-poverty neighborhoods, as part of the Moving To Opportunity (MTO) experiment. This demonstration randomly assigned low-income families living in public housing in five U.S. cities to one of three groups: (a) receipt of a Section 8 housing voucher and help to relocate to a low-poverty area ; (b) receipt of a housing voucher only, with no constraints on the type of destination neighborhood; (c) no relocation.

Kling *et. al.* (2007) and Ludwig *et al.* (2008), among others, have used the randomized housing voucher allocation associated with MTO to examine the impact of relocation to neighborhoods with much lower poverty rates on a very wide set of individual behavioral outcomes including health, labor market activity, crime, education, and more. They find positive effects of the relocation on a variety of health outcomes, including notably mental health, but no effect on education and labor market outcomes.

It is important to note that there are important limitations in the extent to which the treatment effects identified through relocation are informative about the nature of general forms of neighborhood effects per se. First, individuals studied must be eligible for a relocation program in the first place. This typically implies that the resulting sample is special (i.e. so as to be a resident in public housing) and may not be as sensitive to neighborhood effects as other individuals. More generally, even if the eligible population is representative of the target population, the results of an experiment based on a small sample may not extend to broader populations because of the strong possibility that general equilibrium effects may arise in that case.

Strictly speaking, for such experiments to be informative, one needs to assume no interference between units, i.e. that the response of an individual to a treatment is the same whether or not other subjects are administered the treatment. When interference is present, the difference between a treatment group mean and a control group mean does not estimate an average treatment effect, but rather the difference between two effects defined on two distinct sub-populations. This result is very significant: a researcher who fails to recognize it could easily infer that a treatment is beneficial when in fact it is universally harmful [Sobel(2006)].

Second, the experimental design involves relocation to new neighborhoods that are, by design, very different from baseline neighborhoods. This implies that the estimated treatment effect measures the impact of relocating to a neighborhood where individuals initially have few social contacts and where the individuals studied may be very different than the average resident of the new neighborhood. Therefore, any treatment effects identified with this design are necessarily a composite of several factors related to significant changes in neighborhoods that are not easily disentangled.

5 Neighborhood Effects in Job Matching

One important instance of social interactions occurs in the context of informal job referrals in the labor market. Here individuals exchange information about job openings, or refer their social contacts to potential employers, thus affecting labor market outcomes of their "neighbors" (be they proximate in a spatial sense, or in the space spanned by individuals' social networks).

The use of personal networks and referrals in the labor market is very widespread. A consensus estimate, coming from studies that span the past three decades and use a variety of data sources both from the U.S. and from other countries, is that at least half of all jobs are typically found through informal contacts rather than through formal search methods. Further, the use of personal contacts has significant implications for the probability of finding a job, wage earnings, and turnover relative to other search methods [Ioannides and Loury (2004)].

The study of informal hiring practices and labor market referrals has been closely related to the study of neighborhood effects. A direct link between the use of personal networks in job search and the presumed presence of neighborhood effects is given by the observation that social networks have, to some extent, a local dimension in a geographic sense. In a study of Toronto residents in 1978, Wellman (1996) finds that 42% of yearly contacts in individual networks took place with neighbors that lived less than one mile away. Guest and Lee (1983) perform a similar analysis for the city of Seattle, and find that for about 35% of respondents the majority of their non-kin social contacts resided in the same local community. Otani (1999) uses 1986 General Social Survey data for the U.S. (in a comparative Japan-U.S. study) and finds that roughly one in five contacts listed in individual networks are physical neighbors.

Most relevantly to the study of labor market referrals, Lee and Campbell (1999) use data from a 1988 survey of Nashville, Tennessee to look at social ties with immediate neighbors. Their definition of "micro-neighborhoods" consists of "partial face blocks consisting of 10 adjacent housing units each, five on either side of a single street." They find that 31% of these immediate neighbors are judged close or very close by respondents. Further, they specifically ask respondents to whom they would turn for help in finding a job. About 13% of helpers in these networks resided in the respondents' micro-neighborhoods; 73% resided elsewhere in Nashville; the residual 14% were not Nashville residents.

Given this premise, a number of recent studies have looked for evidence that local social interactions within neighborhoods affect employment and wage outcomes. As we discussed earlier, the main estimation problem in the analysis of neighborhood effects (and in the social interactions literature in general) is the possibility that any co-movements in outcomes among members of the same neighborhood (or reference group, more generally) may be due not so much to social interactions but rather to the presence of correlated unobservables at the neighborhood level. Correlation in unobserved attributes may arise because of positive sorting, or because of unobserved shocks that affect the entire neighborhood (for instance, in the case of labor market outcomes, a plant closing that affects employment in the local area), or other unobserved neighborhood-wide institutions that affect the outcome under consideration.

Weinberg *et al.* (2004) use confidential longitudinal data from the NLSY79 to investigate the presence of social interaction effects at the neighborhood level on labor market activity. They also examine the possibility that any correlation in outcomes across neighborhood residents may be explained by the mismatch hypothesis: this theory argues that residents of certain urban neighborhoods incur adverse labor market outcomes because jobs are located far from these neighborhoods. Their identification strategy is to exploit the panel dimension of the data to include individual fixed effects, as well as time-varying individual effects that depend on individual deviations from a typical experience profile.

They find evidence that simple OLS estimation over-estimates the impact of neighborhood social interactions on labor market outcomes, and under-estimates the role of spatial mismatch. They study the impact of employment of adult males in the neighborhood on respondents' annual hours worked. Under their preferred specifications with fixed effects, they find that a one standard deviation increase in neighborhood employment is associated with a 9.5% increase in annual hours worked for adult males on average. In contrast, the effect of a one standard deviation increase in access to jobs (that is, a reduction in spatial mismatch) on hours is a 3.6% increase. When they introduce controls for individual specific experience effects, the estimated effects on hours of neighborhood employment and access to jobs become, respectively, 6.1% and 4.7%.

Topa (2001) analyzes a structural model of transitions into and out of unemployment to estimate the impact of any local social interaction effects on employment outcomes. He posits that individuals may receive useful information about job openings from their employed social contacts but not from their unemployed ones. This positive local feedback implies that the stationary distribution of unemployment in a simulated city exhibits positive spatial correlation. He estimates the model parameters via indirect inference, comparing the simulated spatial distribution of unemployment with the empirical one, using Census data for Chicago in 1980 and 1990.

The identification strategy in that paper relies on: first, the assumption that neighboring census tracts can affect a given tract's employment outcomes only through their employment levels, and not through their own attributes; and second, on the use of local community boundaries (as identified by residents) to distinguish local social interactions from other types of spatially correlated shocks. The results indicate that a one standard deviation increase in neighborhood employment increases expected employment by between 0.6 and 1.3 percentage points. Further, the estimated spillover effects are stronger in tracts with lower education levels and with higher fractions of minorities. This finding is consistent with the direct evidence on referral effects in both sociology and economics.

The analysis also points to an interesting asymmetry: if one raises the amount of information (proxied by neighborhood employment) available in a disadvantaged tract and lowers it in a well-off tract by the same amount, the positive effect on expected employment in the former tract is roughly twice as large, in absolute value, as the negative effect in the latter tract. This is due to the different initial conditions in the two tracts in terms of education levels and other attributes, and the fact that the estimated spillovers vary across these attributes. This has potentially interesting implications for public housing policy, for instance with regard to the idea of dispersing public housing instead of concentrating it in a few areas.

The recent study that makes the strongest case to date for the effects of geographical proximity on job market outcomes is Bayer *et al.* (2008). They adopt a novel strategy to identify and estimate local referral effects in the labor market. The basic idea in this paper is to look for spatial clustering of individual *work* locations for a given *residential* location, as evidence of local referral effects. In order to identify labor market referrals from other spatially correlated effects, the authors estimate the propensity to work together (in the same city block) for pairs of workers who reside in a given city block, controlling for the baseline propensity to work together for residents in nearby blocks (within a reference group of blocks).

The crucial identifying assumption here is that workers can choose residential locations down to a group of blocks, but cannot pinpoint an exact block of choice. Therefore, after controlling for reference group level sorting, the assignment of individuals to specific blocks is essentially random and provides a useful source of variation to identify local referral effects. Measures of the extent of sorting on observable attributes at the block level suggest that this assumption is at least plausible. This paper also employs individual fixed effects to further control for unobserved heterogeneity and examines alternative specifications to address the possibility of reverse causation (i.e., work colleagues giving referrals about desirable residential locations).

Bayer *et al.* (2008) find that, on average, a one standard deviation increase in potential referrals raises expected labor earnings by between 2.0 and 3.7 percentage points, and hours worked per week by between 0.3 and 1.8 hours for men. For women, the effect on labor force participation and expected employment varies between 0.8 and 3.6 percentage points across specifications. Interestingly, they do not find a significant earnings effect for women, which is consistent with the direct evidence on the productivity and wage effects of informal job contacts for women relative to men. Further, the estimated referral effects are stronger for less educated workers, younger workers, and Asians or Hispanics. These results are broadly consistent with direct survey evidence on labor market referrals.

A handful of papers exploit non-neighborhood natural experiments to assess the impact of informal contacts on labor market outcomes. Lalive (2003) studies the possibility that unemployment outcomes may be affected by social interactions among workers by exploiting a "natural" experiment that extended unemployment benefits for a well-specified subset of workers in Austria. The idea is to see whether this exogenous shock to unemployment of eligible workers spilled over to ineligible workers that were in close contact (in terms of their social distance) with a sizable number of eligible workers.

The author essentially uses a difference-in-difference approach that compares the difference in unemployment outcomes for ineligible workers that had a majority vs. a minority of eligible workers in their peer group, across both treatment and control geographic regions. The paper finds significant social interactions effects in unemployment: when the fraction of program eligible workers in one's peer group goes from zero to 100% the risk of long-term unemployment for an ineligible worker increases by 6.7 percentage points. The results appear to be robust to a wide variety of controls for potential general equilibrium effects, differences in local market interactions and social interactions across regions, and unobserved differences in productivity.

Laschever (2008) exploits the random assignment of young American men to the military during World War I to define exogenously constructed peer groups. He then measures the impact of a group's unemployment rate from the 1930 Census on a veteran's own likelihood of being employed. The magnitude of the effect is quite large: a one percentage point increase in his peers' unemployment rate is associated with a half percentage point decrease in one's own expected employment. He further decomposes this effect into the endogenous and the contextual components and finds that the endogenous effect is at least four times as large as the contextual one. This lends some support to the hypothesis that the estimated social effect is due to referrals or informal job contacts.

We also wish to mention two papers that, in our opinion, push the boundaries of current research on neighborhood effects, even though in the context of applications other than job market matching. Cohen-Cole and Zanella (2008) study how usage of public assistance is influenced by one's neighbors behavior. Their study builds on the earlier work by Bertrand *et al.* (2001) by trying to disentangle an information exchange effect from preference interdependence. Assuming one can correctly identify social interaction effects, do individual outcomes comove because individuals share information about eligibility, application procedures and other bureaucratic details, or rather because higher welfare uptake in one's community reduces the stigma associated with such outcomes? The authors' identifying assumption is that information is shared only within racial and ethnic groups, whereas the stigma effect operates across other groups as well. In other words, whereas *social* proximity matters for information exchanges, *spatial* proximity is more relevant for stigma. While this may be an extreme assumption, it seems like a promising first step to disentangle specific mechanisms through which social interactions effects operate.

Conley and Udry (2008) study social learning in technology adoption in the context of the adoption of pineapple cultivation in Ghana. The challenge, again, is how to convincingly distinguish social learning from a situation in which agents' actions comove merely because of common unobservable shocks or attributes. The authors have painstakingly collected direct measures of individual information neighborhoods by tracing individual farmers' networks. They then argue convincingly that these are distinct from geographic neighborhoods that characterize the extent of common growing conditions. Further, they gain a reasonable identifying assumption by analyzing a model of rational learning that delivers the following implication: a farmer's choice of fertilizer will imitate only those of "successful" neighbors. On the other hand, if growing conditions are serially and spatially correlated, the econometrician will observe similar comovements in technology choice regardless of whether neighbors' profits are unexpectedly high or low following adoption.

6 An Agenda

The analysis in Bayer *et al.* (2008) points to the potentially large benefits of analyzing detailed restricted access data available from the Census Research Data Centers, perhaps using a combination of data sources, such as the decennial Census as well as the American Housing Survey data, which are collected at a higher frequency. The innovative identification strategy employed in this paper would not be feasible without the detailed geographic information available in these restricted access data. In addition, it seems particularly promising to expand the set of traditional neighborhood effects questions to include such applications as personal bankruptcy, home foreclosures, and the dynamics of housing supply and demand. In this context, valuable insights will be gained by ongoing efforts by researchers with diverse interests to merge detailed Census data with consumer credit, housing transactions, and mortgage payments panel data.

Several recent works emphasize how much we stand to learn from massive new data sets that have become recently available and are only slowly being utilized. As Kleinberg (2008) points out, the past decade has witnessed a coming-together of the technological networks that connect computers in the internet and the social networks that have always connected people since the emergence of human societies. Research by Mayer and Puller (2007), who merge confidential administrative data and information from `Facebook.com` to study the structure and composition of social networks on university campuses and investigate the processes that lead to their formation, is indicative of the potential in this area.

Gonzalez *et al.* (2008) is a good case in point on what we can learn from these vast new data sources. They use data from 100,000 anonymous mobile phone users whose positions are tracked over a six-month period. They show that human trajectories show high temporal and spatial regularity, with time-independent individual time paths and regular frequenting of particular locations. Once corrected for individual characteristics, paths obey a single spatial distribution. One would think that this is likely to be influenced by physical infrastructure of the locales where they live. This appears to be in contrast to previous research that suggests the human movements follow random walks with fat tailed displacements and waiting-time distributions. Data of this type go much beyond the naive but common perception that what can be learned about social interactions would be based only on social networking sites and the like. Modern information and communication technologies mediate numerous modes of interactions that may be directly interpersonal or involve self-expression in the intellectual, scientific or opinion sphere.

We also believe it is very important to go beyond the simple detection and estimation of generic social interaction effects to try to better understand and disentangle the specific mechanisms through which social effects operate. For instance, is it social learning or social influence? In other words, is it information or preferences? One possible approach to identify these separately consists of using different assumptions about the shape of these mechanisms as a function of the number of social contacts. There is some evidence [see Bandiera and Rasul (2006)] that an information effect may be concave in number of contacts. This makes intuitive sense as it is easy to model an information exchange process in which the marginal "new" information acquired from additional contacts is less and less valuable. On the other hand, a social influence effect might very plausibly be linear or even convex. For instance, the reduction in stigma associated with applying for public assistance could grow proportionately larger as more and more people use welfare. Work in progress by Bayer, Ross, and Topa aims at exploring this approach using different functional form assumptions for various effects.

Another important area for future research concerns the endogeneity of social networks: how do networks change over time, as agents rationally anticipate the effect of establishing/severing a tie on their future payoffs? Bisin *et al* (2006) analyze a model of rational social interactions, in which agents are forward-looking and evaluate future costs and benefits of social ties. This seems like a promising approach that may deliver useful testable implications for the shape of social interaction effects and help identify them separately from correlated effects. On the empirical front, several working papers (Bifulco *et al.* (2008), Weinberg (2007)) use detailed data on individual networks and outcomes of teenagers (AddHealth), to estimate models of endogenous network formation and evolution over time and the impact on behavior [Fryer and Torelli (2006)].

On the theoretical front, three excellent new books on networks, Goyal (2007), Jackson (2008) and Vega-Redondo (2007), will likely help integrate the large amount of theoretical research on network theory. New advanced on econometric techniques, like by Bramoullé *et al.* (2009) and others, offer new ways to take advantage of network structure in identifying decisions in network settings. They can be augmented by means of non-parametric techniques, like the ones proposed by Brock and Durlauf (2007), who examine the identifiability of parameters of binary choice with social interactions models when the distribution of random payoff terms is unknown. Their results on partial identification of endogenous social interactions in some special cases of *pattern reversals* (between contextual effects and endogenous outcomes) are particularly interesting and lend themselves to network-based extensions.

More generally, it would be very useful to write general models in order to derive implications for the joint distribution of income, human capital, possibly ethnicity by integrating job network effects on the one hand with sorting and intergenerational transmission processes on the other. While reminiscent of ideas in Benabou (1993), this would be a novel research agenda and would involve writing down processes at different time scales (business cycle frequency, life cycle frequency, and intergenerational frequency) and then solving for their equilibrium invariant distribution. It would then be interesting to think about how individual networks might evolve at these different time scales. At the cyclical frequency, people might optimally adjust their network as well as the number of contacts (which in some sense is a measure of search effort) depending on the value of search and their cost of time. At the same time, people may adjust their network size and/or composition based on the previous generation's experience (thus providing another inter-generational linkage). Location–specific social networks can provide additional motivation for community-based amenities, that may reflect peer effects [Epple *et al.* (2009)].

In addition, it would be fruitful to incorporate the use of personal contacts into full-fledged macro models of job search a la Mortensen and Pissarides, Shimer, etc. This may give us useful insights into the ways in which an aggregate matching function may change as a function of alternative search methods and intensities. Further, the choice of search method could be endogenized in the context of a formal search model, in order to study how the choice between formal and informal methods may vary according to different market conditions, and what implications that has on outcomes. Empirically, it would also be useful to study whether the use of personal contacts varies with the business cycle.

Finally, it would be very interesting to incorporate informal job search methods into urban economics models, to see how they affect neighborhood dynamics, as well as equilibrium wage and rent distributions within a city. Zenou (2008) takes a first stab at this research agenda, by analyzing the role of informal contacts and referrals in an explicit urban model in which agents are located at varying distances from jobs, and space affects the extent of social interactions. He finds that an increase in weak tie interactions lowers unemployment but also increases equilibrium rents in the city. Under certain conditions, an increase in weak ties also induces a higher equilibrium wage.

A number of areas of neighborhood effects research are particularly relevant for policy. Moffitt (2001) lays down a basic approach. Large-scale social experiments, like MTO, as well as small-scale ones, like METCO, naturally lend themselves to formal evaluation by means of the tools of the body of knowledge known as neighborhood effects. They may serve as settings for assessing arguments in favor and against deliberate racial and income integration of, respectively, residential communities and of public educational institutions as forms of social policy.

Alongside more traditional, large-scale randomized experiments, other, small-scale social experiments in local neighborhoods, social networks and peer effects may also be quite fruitful. Falk and Ichino (2006) conduct a field experiment in which worker groups are randomly formed, and find that worker productivity is affected by peers in one's team; Palacios-Huerta (2003) runs social interactions experiments in the lab and finds that learning contains a social interactions dimension.

7 References

Angrist, Joshua D., and Kevin Lang. 2004. "Does School Integration Generate Peer effects? Evidence from Boston's METCO Program," *The American Economic Review,* 98, 1613–1634.

Arcidiacono, Peter, and Sean Nicholson. 2005. "Peer Effects in Medical School, " *Journal of Public Economics,* 89, 327–350.

Bandiera, Oriana and Imran Rasul. 2006. "Social Networks and the Adoption of New Technology in Northern Mozambique", *Economic Journal,* 116, 862–902.

Bayer, Patrick, and Stephen Ross. 2008. "Identifying Individual and Group Effects in the Presence of Sorting: A Neighborhood Effects Application," working paper, Department of Economics, University of Connecticut, and Department of Economics, Duke University, March.

Bayer, Patrick, Stephen Ross and Giorgio Topa. 2008. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," *Journal of Political Economy,* 116 (6), 1150–1196.

Benabou, Roland A. 1993. "Workings of a City: Location, Education, and Production," *Quarterly Journal of Economics,* 108, pp. 619–652.

Bertrand, Marianne, Erzo F. P. Luttmer, and Sendhil Mullainathan. 2000. "Network Effects and Welfare Cultures," *Quarterly Journal of Economics*, 115(3), pp. 1019–1055.

Bifulco Robert, Jason Fletcher and Stephen Ross. 2008. "The Effect of Classmate Characteristics on Individual Outcomes: Evidence from the Add Health", working paper, University of Connecticut, Department of Economics.

Bisin, Alberto, Andrea Moro and Giorgio Topa. 2009. "Peer Effects vs Personal Preferences in Smoking Choices: Estimating a Social Interactions Model", working paper, Department of Economics, Vanderbilt University.

Bisin Alberto, Ulrich Horst, Onur Özgür. 2006. "Rational Expectations Equilibria of Economies with Local Interactions", *Journal of Economic Theory,* 127(1), 74–116.

Bramoullé, Yann, Habiba Djebbari, Bernard Fortin. 2009. "Identification of Peer Effects through Social Networks," *Journal of Econometrics,* forthcoming.

Brock, William A., and Steven N. Durlauf. 2001. "Interaction-Based Models." Heckman, James, and Edward Leamer, eds. *Handbook of Econometrics, Vol. 5.* North-Holland, Amsterdam, 3297–3380.

Brock, William A., and Steven N. Durlauf. 2006. "Multinomial Choice with Social Interactions," in Blume, Laurence E., and Steven N. Durlauf. (eds.) *The Economy as an Evolving Complex System, III.* Oxford University Press: Oxford and New York, pp. 175–206.

Brock, William A., and Steven N. Durlauf. 2007. "Identification of Binary Choice Models with Social Interactions," *Journal of Econometrics,* 140, 52–75.

Clampet-Lundquist, Susan and Douglas S. Massey. 2008. "Neighborhood Effects on Economic Self-sufficiency: A reconsideration of the Moving to Opportunity Experiment," *American Journal of Sociology,* 114(1), 107–143.

Cohen-Cole, Ethan, and Jason Fletcher. 2008. "Is Obesity Contagious? Social Networks versus Environmental Factors in the Obesity Epidemic," *Journal of Health Economics,* 27, 1382–1387.

Cohen-Cole, Ethan, and Giulio Zanella. 2008. "Welfare Stigma or Information Sharing? Decomposing Social Effects in Social Benefit Use," working paper, Department of Economics, University of Siena, November.

Conley, Timothy G., and Christopher R. Udry. 2008. "Learning About a New Technology: Pineapple in Ghana," *American Economic Review,* forthcoming.

Datcher Linda. 1982. "Effects of Community and Family Background on Achievement." *The Review of Economics and Statistics,* 64, 32–41.

Drewianka, Scott. 2003. "Estimating Social Effects in Matching Markets: Externalities in Spousal Search." *Review of Economics and Statistics,* 85, 408–423.

Durlauf, Steven N. 2004. "Neighborhood Effects." in Henderson J. Vernon, and Jacques-Francois Thisse, (eds.) *Handbook of Urban and Regional Economics, Vol. 4, Cities and Geography,* Amsterdam: North-Holland, pp. 2173–2242.

Epple, Dennis, Brett Gordon, and Holger Sieg. 2009. "Drs. Muth and Mills Meet Dr. Tiebout: Integrating Location-Specific Amenities into Multi-Community Equilibrium Models," paper presented at the "Journal of Regional Science 50th Anniversary Conference," Federal Reserve Bank of New York, New York.

Falk, Armin, and Andrea Ichino. 2006. "Clean Evidence on Peer Effects," *Journal of Labor Economics,* 24(1), 39-57.

Fryer, Roland G., and Paul Torelli. 2006. "An Empirical Analysis of 'Acting White.'" working paper, Department of Economics, Harvard University, October; NBER Working Paper No. 11334.

Glaeser, Edward L., Bruce I. Sacerdote and Jose A. Scheinkman. 2003. "The Social Multiplier," *Journal of the European Economic Association,* 1, 345–353.

González, Marta, César A. Hidalgo, and Albert-László Barabási. 2008. "Understanding Individual Human Mobility Patterns," *Nature* 453(5), pp. 778–782.

Goyal, Sanjeev. 2008. *Connections: An Introduction to the Economics of Networks.* Princeton, NJ: Princeton University Press.

Graham, Bryan S. 2008. "Identifying Social Interactions through Conditional Variance Restrictions," *Econometrica,* 6(3), 643–660.

Guest, Avery M. and Barrett A. Lee. 1983. "The Social Organization of Local Areas," *Urban Affairs Quarterly,* 19, 217–240.

Ioannides, Yannis M., and Linda D. Loury. 2004. "Job Information Networks, Neighborhood Effects, and Inequality," *Journal of Economic Literature*, 42, 1056–1093.

Ioannides, Yannis M., and Jeffrey E. Zabel. 2008. “Interactions, Neighborhood Selection, and Housing Demand.” *Journal of Urban Economics,* 63, 229–252.

Jackson, Matthew O. 2008. *Social and Economic Networks.* Princeton, NJ: Princeton University Press.

Kling, Jeffrey R., Jeffrey B. Liebman, and Laurence F. Katz. 2007. “Experimental Analysis of Neighborhood Effects.” *Econometrica,* 75(1), 83–119.

Lalive, Rafael. 2003.“Social Interactions in Unemployment”, discussion paper no.803, Institut zur Zukunft der Arbeit, `http://www.iza.org/`.

Laschever, Ron. 2008. “The Doughboys Network: Social Interactions and the Employment of World War I Veterans”, working paper, Department of Economics, University of Illinois, Urbana–Champaign, March.

Lee, Barrett A. and Karen E. Campbell. 1999. “Neighbor Networks of Black and White Americans.” , in Barry Wellman, (ed.),*Networks in the Global Village: Life in Contemporary Communities*, Boulder, Colorado: Westview Press, pp. 119–146.

Ludvig, Jens, Jeffrey R. Kling, Laurence F. Katz, Jeffrey B. Liebman, Greg J. Duncan, and Ronald C. Kessler. 2008. “What Can We Learn about Neighborhood Effects from the Moving to Opportunity Program,” *American Journal of Sociology,* 114(1), 144–188.

Manski, Charles F. 1993. “Identification of Endogenous Social Effects: The Reflection Problem,” *Review of Economic Studies,* 60, 531–542.

Mayer, Adalbert, and Steven L. Puller. 2007. “The Old Boy (and Girl) Network: Social Network Formation on University Campuses,” *Journal of Public Economics,* 92, 329–347.

Moffitt, Robert A. 2001. “Policy Interventions, Low-Level Equilibria and Social Interactions,” in Durlauf, Steven N., and H. Peyton Young (eds.), *Social Dynamics,* Cambridge, MA: The MIT Press, pp. 45–82.

Otani, Shinsuke. 1999. "Personal Community Networks in Contemporary Japan," in Barry Wellman (ed.), *Networks in the Global Village: Life in Contemporary Communities*, Boulder, Colorado: Westview Press, pp. 279–297.

Palacios-Huerta, Ignacio. 2003. "Learning to Open Monty Hall's Doors," *Experimental Economics,* 6(3), 235-251.

Sacerdote, Bruce. 2001. "Peer Effects With Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics*, 116, 681–704.

Sirakaya, Sibel. 2006. "Recidivism and Social Interactions," *Journal of the American Statistical Association,* 101(475), 863–875.

Sobel, Michael. 2006. "Spatial Concentration and Social Stratification: Does the Clustering of Disadvantage 'Beget' Bad Outcomes?" Ch. 8, in Bowles, Samuel, Steven N. Durlauf, and Karla Hoff (eds.) *Poverty Traps.* Princeton, NJ: Princeton University Press, pp. 204–229.

Topa, Giorgio. 2001. "Social Interactions, Local Spillovers and Unemployment," *The Review of Economic Studies*, 68(2), 261–295.

Vega-Redondo, Fernando. 2007. *Complex Social Networks,* Cambridge, UK: Cambridge University Press.

Weinberg, Bruce A. 2007. "Social Interactions with Endogenous Associations," National Bureau of Economic Research, Working Paper 13038, April.

Weinberg, Bruce A., Patricia B. Reagan, and Jeffrey J. Yankow. 2004. "Do Neighborhoods Affect Hours Worked? Evidence from Longitudinal Data," *Journal of Labor Economics,* 22(4), 891–924.

Wellman, Barry. 1996. "Are Personal Communities Local? A Dumptarian Reconsideration." *Social Networks,* 18, 347–354.

Zenou, Yves. 2008. "Social Interactions and Labor Market Outcomes in Cities", working paper, Department of Economics, Stockholm University.