CHAPTER 25

Neighborhood Effects And Housing

Yannis M. Ioannides[*,†]

Contents

[*] Department of Economics, Tufts University, Medford, MA 02155, USA. 617 627 3294. yannis.ioannides@tufts.edu
[†] I am grateful to Giorgio Topa for very careful comments on an earlier draft and to Steven Durlauf, Anna Hardman, Kurt Schmidheiny and Jeff Zabel for their very generous comments on related work presented in earlier papers. I am solely responsible for any errors.

Handbook of Social Economics, Volume 1B
ISSN 0169-7218, DOI: 10.1016/S0169-7218(11)01030-6

Abstract

This chapter focuses on neighborhood effects in housing markets. Households in effect choose neighborhood effects, or more generally social interactions, via their location decisions, which renders them endogenous. Across several classes of models that it examines, it emphasizes how we may detect empirically the presence of neighborhood effects when they may be priced by housing markets and be capitalized into housing values and rents. The chapter focuses on models that are empirically relevant and help identify neighborhood effects, and discusses actual empirical findings.

The first class of models examined involves models of choice over discrete sets of individual dwelling units that allow for a multidimensional bundle of characteristics. These models extend the Berry–Levisohn–Pakes characteristics-based approach and allow for endogenous contextual effects. The chapter develops in detail a specific application that also endogenizes contextual effects at a low level of dimensionality. The chapter examines neighborhood choice, with endogenous and contextual neighborhood effects, and housing demand (with housing being measured as a scalar) as joint decisions. This approach utilizes individual and neighborhood-level data at several levels of aggregation.

The chapter next examines neighborhood effects within the canonical Alonso–Mills–Muth urban model with a well-defined spatial structure. Finally, the chapter reviews hierarchical models of neighborhood location in the presence of social interactions. These models describe communities in terms of a low-dimensional vector of attributes that are aggregated into a public good whose consumption is nonrival. Such approaches are designed to utilize community-level data, along with information on the community-specific distributions of various sociodemographicsociodemographic characteristics of individuals.

JEL Codes: A10, C00, C31, C35, C81, D01, R14, R20, R21, H40, Z10, Z13

Keywords

Neighborhood effects
housing
endogenous social interactions
neighborhood choice
hedonics
urban spatial structure

1. INTRODUCTION

Individuals and firms regularly face location decisions that require them to assess a multitude of factors. Households who decide where to live within urban areas consider access to transportation routes, to schools, and to various local and citywide amenities. Local amenities range from attractive buildings and natural scenery (in the form of parks and natural settings), on one hand, to characteristics (and even habits) of other residents close by and to activities in which those other individuals engage. Downtown areas of vibrant cities are often acting as magnets for a variety of formal and informal activities. Music, theater and all forms of cultural and entertainment activities as well as public buildings that themselves connote history and culture confer character to city centers throughout the world. Occasionally, however, changing fads shift attention to other areas within a single large metropolis. In fact, different areas may coexist offering mixes of diverse activities, and at the end, conferring character to entire metropolitan areas, as well.

Some of these activities are so persistent that they come to signal elements of specialization for particular locales within large cities. The West End and the City in London, the Left Bank in Paris, downtown New York, Harvard Square in Cambridge, Massachusetts, Hollywood and Santa Monica in Los Angeles, and so on are cases in point. This can also be said about large cities and metropolitan areas throughout the world, and regardless of the state of development. Why is it that urban activities sometimes conspire to produce true world landmarks and evoke of great things and other times result in failures? [Dougherty (2002)] All such urban features are manmade. The theories reviewed in this chapter take the view that they are all driven by individuals' quest for being near other individuals, either because of who those others are or of what they do. In a nutshell, they are driven by neighborhood effects.

The recent explosion of interest in the role of social interactions calls for a framework for thinking about these questions. Specifically, it is important to be able to distinguish between different effects within individuals' social milieus. Being guided by the canonical framework proposed by Charles Manski [Manski (1993; 2000)], in particular, we wish to distinguish between different types of social effects. Consider first that individuals may value economic decisions, or actions, more generally, by members of their reference groups. For example, my neighbors remodel their house, or simply keep up its maintenance in ways that shame me, if I do not keep up. My children's hearing of academic, athletic and other accomplishments of other children in the neighborhood motivates them to imitate them or even to react in a nonconformist way. These types of effects are known as *endogenous* social effects, because they originate in deliberate decisions by members of one's reference groups. Groups of people who regularly interact among themselves, form, often by choice, reference groups.

Individuals may value the actual characteristics of others in their social and residential milieus and deliberately seek to choose among alternative sets of such characteristics. Effects that originate in the characteristics of members of one's reference groups are known as *exogenous, or contextual* effects, and are, of course, *social* effects. When different individuals tend to act similarly because they possess similar characteristics or face similar institutional environments, we say that they are subject to *correlated* effects. Naturally, all these effects may coexist, posing challenging problems when one wishes to distinguish econometrically among them.[1] There may be very good reasons, including for the purpose of policy design, to want to distinguish among them.

For example, how can policy affect social outcomes at the urban neighborhood or community level? Specifically, in the context of the urban economy, could we engineer improvements in living conditions for the residents of "depressed" or disadvantaged areas, by encouraging the relocation of individuals with particular characteristics? Two actual policy options are worth contemplating in this context, and operate on the supply and demand, respectively. Increasing the supply of affordable housing within otherwise high-housing cost communities is one such policy option. Residents who value proximity to demographically more diverse groups would be better off, but others might be worse off. The net effect depends on neighborhood effects. Another policy is subsidizing the relocation of low-income households out of disadvantaged and into more prosperous communities. Again, neighborhood effects in the form of role models may confer benefits on relocated households, but they may impact adversely incumbent households. Related conceptually are such important matters as what determines the character of urban communities, their ambience. Whether policy may affect urban ambience depends on neighborhood effects. If such effects are absent, those policies may have effects only through prices.

Individuals decide about joining clubs (or gated communities, or other types of residential communities with controlled access, like New York City co-ops) in order to avail themselves of the associated services being offered. But individuals also value clubs precisely because they allow them greater choice over the types of other individuals they are likely to interact with. Individuals sort themselves across such voluntary associations. Individuals choosing public communities and neighborhoods avail themselves of access to community-based and typically shared amenities. The fact that many amenities of communities are open to the public, regardless of residence, implies a lesser degree of control over who one would likely come into

[1] See Durlauf and Ioannides (2010) for a review of the state-of-the-art on the estimation of social interactions models and Blume, Brock, Durlauf and Ioannides (2010) for an exhaustive presentation of the state-of-the-art on identification of social interactions.

contact with. However, the cost of housing and community-based taxes serve as (indirect) admission prices.

This discussion helps clarify the broader issues associated with community formation. Availability of excludable services and access to nonexcludable local public goods combine to define attractiveness of a community. New residents enter, as long as expected attractiveness dominates other options. Since housing prices help ration access to public communities, it follows that market prices themselves reflect, and therefore also proxy for, the set of attributes that characterize communities, including social effects. Precisely because individuals take the price of a good as given and beyond their control, and make their decisions according to the enjoyment they expect to derive from different alternatives, equilibrium prices aggregate information and operate as signals of the characteristics of all market participants. It is also information about nonparticipants as well, to the extent that some individuals may be priced out of a particular market that help define its bounds.

In deciding whether or not to locate in a particular city or neighborhood, individuals weigh all relevant factors, such as market variables, like prices and rents, contextual variables and endogenous effects. They do so from their own perspectives and by forming expectations about what these effects are likely to be. In other words, populations of heterogenous individuals sort themselves into different communities. When they do locate by pursuing optimizing strategies, their individual characteristics once they have been aggregated contribute to defining prices and the distributions of characteristics by location. Thus, they help define a social equilibrium for a set of communities. It is in this sense that housing prices are hedonic prices, a concept that we will explore further in the remainder of this chapter.

Some of the sorting we observe is of course sorting on directly, or even indirectly, observable information. An example of the latter is that families might seek information on the quality of publicly provided education, recreation, and other such amenities, which is often widely available in the US context, before deciding where to locate. This is the focus of much of the research on hedonic prices whose pitfalls are better understood due to extensive research that has followed the influential work of Rosen (1974).

As Rosen (2002) underscores, it is important to assess such sorting in order to, inter alia, understand the social valuation of neighborhood amenities when individuals differ in terms of preferences. For example, if neighborhood safety is valued differently by different people, those who value it less sort to less safe neighborhoods. Estimating the average value of neighborhood safety to society based on valuations by those who sort to less safe neighborhoods would bias it downwards, and vice versa for those who sort to safer neighborhoods. With sorting, we are likely to see different types of neighborhoods differing with respect to prevailing neighborhood safety. In practice, how do we know what particular factors are most important in individuals' decisions?

And what if some of those are unobservable, or involve social interactions? Sorting that rests on unobservables makes it difficult to determine which factors are responsible for sustaining neighborhoods and communities that differ significantly in terms of attributes. There is an inherent difficulty in distinguishing among alternative factors that drive individuals' location decisions. Are they due to attraction of local synergies (or anticipated spillovers from locational decisions of individuals) or to underlying and possibly only partly unobservable natural advantages? In either case, these factors work to bring and hold people together.

This chapter elaborates further on the role of social interactions in location decisions by individuals by working with both aspatial and spatial models. The latter type of models result by placing a location model within an urban space setting with a well defined geographic metric, like distance from a city center. The former are akin to choosing membership to a social group from among a number of alternative social groups. Whereas clubs may exclude at will individuals from becoming members, access to communities is subject to free entry and is rationed by market-determined housing prices and taxes, which are set by local governments. This also suggests an element of similarity between social interactions and local public goods.

General multinomial choice models, as for example the one invoked by Brock and Durlauf (2002; 2007), are readily capable of handling aspatial location decisions. Some of the location decisions examined here are quite closely related to group choice. Such a general framework allows one to study multidimensional attributes of a decision, like the size of a dwelling unit and its different characteristics, including its location. It is natural to think that several attributes of a location decision emanate from a single optimization problem. For example, location is a discrete choice from among alternative sets of opportunities (such as communities), within each of which individuals may select continuous quantities of interest. Of course, such decision structures are testable, in principle.

While this chapter aims at understanding the choice of housing in the presence of neighborhood effects, the approach applies equally well to choice in labor markets, and to joint housing and job location. Adapting a stochastic location model to an urban setting allows us to evaluate the role of proximity to urban centers—the hallmark of urban economics—and of the characteristics of one's neighbors when all location decisions are endogenous.

The first class of models we examine in Section 2.1 involves models of choice over discrete sets of individual dwelling units that allow for multidimensional bundles of characteristics. Models of these types, which borrow from the industrial organization literature and especially the Berry–Levisohn–Pakes characteristics-based models [Berry (1994); Berry, Levinsohn and Pakes (1995; 2004)], have been developed by Bayer and a number of coauthors. They lead naturally to hedonic models while sorting is accounted for. In the context of the neighborhood effects literature, these models allow for endogenous contextual effects.

We pursue further in Section 2.5 a conceptually related approach, due to Nesheim, that also endogenizes contextual effects but involves much lower dimensionality. That is, individuals choose neighborhoods while recognizing that their neighbors' characteristics along with their own determine educational outcomes for their children. This approach allows us to obtain in closed form equilibrium housing price functions that are consistent with hedonic valuation of neighborhood attributes.

The paper turns next to aspatial models of neighborhood choice, with endogenous and contextual neighborhood effects, and housing demand (with housing measured as a scalar) as joint decisions. It emphasizes a model, due to Ioannides and Zabel (2008), which is designed to utilize individual and neighborhood-level data at several levels of aggregation. This approach links naturally with hierarchical choice models. Hierarchical models of neighborhood location, which originate in Epple and Sieg (1999) and have been developed further by Epple and several other co-authors, describe communities in terms of a low-dimensional vector of attributes that may be aggregated into a public good. Individuals' choice of community is subject to community-specific housing price and tax rate, which at equilibrium must sustain individuals' choices. This approach is designed to utilize data that are aggregated at the community-level and to match them with the community-specific distributions of various sociodemographic characteristics of individuals. When considered against hedonic-type models, the housing price again does double duty, by pricing housing and "admission" into communities (or neighborhoods). In these models, sorting across communities works either only through individuals' valuation of community-specific amenities or (in their most recent versions) with community-specific amenities combined with neighborhood effects in the form of contextual effects.

The chapter also takes up neighborhood effects within the canonical Alonso–Mills–Muth urban model with a well-defined spatial structure and individuals commuting to a predetermined central business district. If individuals differ with respect to income or preference characteristics, then the standard urban model implies segregation. The paper discusses extensions of the model that allow for amenities that may be exogenous or endogenous and are spatially dispersed, and naturally influence individuals' location decisions and the associated housing price structure.

2. SPATIAL MODELS OF LOCATION WITH SOCIAL INTERACTIONS

The models and empirical results that we present first in effect treat symmetrically the entire set of factors that determine location decisions. We start with a specific model that originates in a broader line of research that emphasizes social interactions that have consequences for individuals' decisions, in general as well as in particular in the context of housing decisions. We refer to such models as aspatial because they do not aim at explaining the spatial organization of urban neighborhoods nor urban geometry as such, and therefore do not make distance from the central business district (CBD) a key element of the analysis.

Location decisions are interesting in the context of social interactions because proximity to others is a key attribute of housing and enters housing decisions, more generally. Neighbors typically know one another and may acquire intimate knowledge of each others' habits and indeed lives. The composition of urban neighborhoods in terms of the socioeconomic characteristics of their residents and the mix of economic and social activities is an important element of their attractiveness [*c.f.* Becker and Murphy (2000)]. Different communities differ in terms of density of settlement, quality of amenities and public services, distance to employment centers and to transportation systems, to name just a few. Our analysis of aspatial models probes our understanding of what brings and holds people together.

2.1 Models of choice among individual dwelling units

In the models we discuss here, different dwelling units are seen as differentiated commodities, with differentiation in a multitude of dimensions. The basic model in this class of papers is essentially an application of the differentiated products approach, typically being referred to as the *BLP approach* [Berry (1994); Berry, Levinsohn and Pakes (1995)], to housing markets. It is a natural approach to housing decisions, precisely because dwelling units are hardly a standardized commodity.

In a number of papers by Bayer *et al.*, households choose among individual dwelling units taking into consideration exogenous characteristics of neighborhoods in which units lie as well the characteristics of neighbors, in other words, contextual effects. We start from the basic model, as presented in Bayer, Ferreira and McMillan (2007a, 2007b) and Bayer, McMillan, and Rueben (2009). Individual i, $i \in \mathcal{I}$, values a dwelling unit h in terms of a vector, $\mathbf{X}_h$, of characteristics which includes dwelling size, age, type, tenure status, and neighborhood characteristics such as crime, school quality, socioeconomic composition of neighborhood and geography. Dwelling unit h carries a price p_h. Household i chooses its residence from among a discrete choice set of options, $h \in \mathcal{H}^i$, so as to maximize

$$V_h^i = \alpha_X^i \mathbf{X}_h - \alpha_p^i p_h + \xi_h + \varepsilon_h^i, h \in \mathcal{H}^i, \tag{1}$$

where the random variable ξ_h is specific to dwelling unit h, and thus common to all households that consider that unit, and captures the unobserved quality of the unit and its neighborhood; and ε_h^i denotes a random variable that household i draws from a specified distribution. Associated with the definition of the choice set is the appropriate price p_h that enter the choice probabilities.

It is an important component of the BLP approach that the set of magnitudes α^i_X, α^i_p be specified as functions of individual characteristics as follows:

$$\alpha^i_j = \alpha_{0j} + \sum_{r=1}^{R} \alpha_{rj} Z^i_r, j \in \{X, p\}, \tag{2}$$

where r indexes the components of the vector of observable characteristics $\mathbf{Z}^i$, that is individual i's own socioeconomic characteristics.[2] With this specification, it is possible to rewrite V^i_h from (1) so as to distinguish a unit h–specific term, mean indirect utility δ_h,

$$\delta_h = \alpha_{0X}\mathbf{X}_h - \alpha_{0p} p_h + \xi_h, \tag{3}$$

from a term that interacts unit-specific with individual-specific variables, λ^i_h,

$$\lambda^i_h = \left(\sum_{r=1}^{R} \alpha_{rX} Z^i_r \right) \mathbf{X}_h - \left(\sum_{r=1}^{R} \alpha_{rp} Z^i_r \right) p_h,$$

and a random shock ε^i_h. That is:

$$V^i_h = \delta_h + \lambda^i_h + \varepsilon^i_h. \tag{4}$$

It is straightforward to obtain the probability of individual i's choice of dwelling unit h from her choice set $h \in \mathcal{H}^i$, once that choice set and the distribution from which the ε^i_h's are drawn have been specified. While it is convenient to assume that the stochastic shock ε^i_h in (4) is extreme-value distributed, as Bayer *et al.* in fact do, in view of Ellickson (1981) it is not necessary to do so.[3] If the ε^i_hs are extreme-value distributed of type II, with mean zero, variance $\frac{\pi^2}{6\varrho^2}$, and mode $-\frac{EC}{\varrho}$, where EC $\approx$ 0.5772, denotes Euler's constant, then the choice probabilities are given by:

$$\text{Prob}^{\,i}_h = \frac{\exp\left[\delta_h + \lambda^i_h\right]}{\sum_{k \in \mathcal{H}^i} \exp\left[\delta_k + \lambda^i_k\right]}, h \in \mathcal{H}^i, \tag{5}$$

where we have normalized by setting $\varrho = 1$ in (5). Given characteristics of dwelling units $\mathbf{X}_h$ and of individuals $\mathbf{Z}^i$, and a specification of the distribution of unobservable characteristic in Equation (3), ξ_h, estimation of the discrete choice model by means of maximum likelihood naturally *forces* the probabilities that each unit would be occupied to sum up to 1 [*op. cit.*, Equation (12)]. The estimation process delivers estimates of the δ_h's, that is, the unit-specific component of the valuation, $\delta_h = \alpha_{0X} X_h - \alpha_{0p} p_h$

[2] The stochastic structure in Berry *et al.* (1995) is even more general.

[3] The theory of extreme order statistics offers conditions under which the extreme value distribution emerges as a limit of a bidding process that underlies the allocation of dwellings to individuals. See Ellickson (1981) and Jaïbi and ten Raa (1998).

$+ \xi_h$, for each of the units in the sample. Note also that all of the data that enter the problem bear upon the estimation of the δ_h's. We may also estimate α_{rX}, α_{rp}. It is then possible, in principle, to estimate parameters (α_{0X}, α_{0p}) and the distribution of the ξ_h's along the lines of (3), by using the estimated unit-specific δ_h's as dependent variables in a second-stage regression.

Although ξ_h itself may be specified stochastically to be independent of the other regressors, the values of the variables that are associated with a particular unit h, X_h, p_h are themselves subject to selection bias. This applies particularly to the price that is appropriate for each unit. It is this key issue that motivates our approach to hedonic theory further below.

Bayer *et al.* recognize this and follow the original BLP literature by instrumenting the price by means of attributes of houses and neighborhoods beyond the immediate neighborhood of where a particular household has chosen to live. Such instruments contain information about the overall housing demand but are assumed to be uncorrelated with the realizations of the specific variables that enter the determination of demand in the immediate vicinity of unit h. As they put it, such an instrumenting strategy exploits an inherent feature of the sorting process — that the overall demand for houses in a particular neighborhood is affected not only by the features of the neighborhood itself, but also by the way these features relate to the broader landscape of houses and neighborhoods in the region.

2.2 Hedonic price indices

The development so far addresses the valuation of different dwelling units by individuals. What does this model imply about the *market valuations* of different dwelling units? A simple way to start is to exclude λ_h^i from the definition of unit-specific utility (4). This forces the estimates of δ_h to be equal for all units, and may therefore be set equal to 0. When the individualized component of utility, λ_h^i, is excluded, then the estimates of δ_h are equal to zero for all units. Equation (3) then yields by solving for the price p_h of a unit h:

$$p_h = \frac{\alpha_{0X}}{\alpha 0p}\mathbf{X}_h + \frac{1}{\alpha 0p}\xi_h. \tag{6}$$

This is like a standard hedonic price regression that relates dwelling unit prices p_h to their observable characteristics $\mathbf{X}_h$ and to unobservable characteristics of their neighborhoods. However, suppression of the individualized component λ_h^i removes contextual effects from the model of housing demand. The upshot of this is that regressors controlling for attributes that are positively correlated with the vector of indirect utilities will have coefficients that will be downwards biased.

There is another noteworthy aspect of hedonic regressions. Units chosen by a particular group of the population will be associated with higher utility due to revealed preference, by an amount given by the term δ_h in (4). By omitting those terms from the estimation would lead to understating the willingness to pay for unit characteristics.

It is possible to obtain a systematic accounting for these effects. We first obtain an expression that summarizes the value to a decision maker associated with the choice process. The expected value $\widetilde{V}^i_{\mathcal{H}}$ of the maximum utility from the choice process for individual i, is given by:

$$\mathcal{E}\left\{\max_{h\in\mathcal{H}^i} : \widetilde{V}^{\mathcal{H}^i}_i(\mathbf{X};\mathbf{Z}^i)\right\} = \ell n\left(\sum_{h\in\mathcal{H}^i} \exp\left[\delta_h + \lambda^i_h\right]\right) = \ell n \int \exp\left[\delta_h + \lambda^i_h\right] \cdot dF_{\mathbf{X}_h}, \qquad (7)$$

where $F_{\mathbf{X}_h}$ denotes the distribution of dwelling unit characteristics. The full distribution, not just the mean, is also available in closed form.[4]

Working in a dual fashion, we define the market valuation of each unit as the outcome of bidding among the set of individuals, $i \in \mathcal{I}$. Under our assumptions, the maximum valuation of a particular unit h by all individuals in the sample, defined as

$$\widetilde{V}^{\mathcal{I}}_h \equiv \max_{i\in\mathcal{I}} : \widetilde{V}^i_h(\mathbf{X};\mathbf{Z}^i),$$

has a probability distribution defined by:

$$\text{Prob}\left[\max_{i\in\mathcal{I}} : \quad \delta_h + \lambda^i_h + \varepsilon^i_h \leq \nu\right] = \exp\left[-e^{-\nu}\sum_{i\in\mathcal{I}} \exp\left[\delta_h + \lambda^i_h\right]\right].$$

This may be simplified by defining $\widetilde{\nu} \equiv \ell n\left[\sum_{i\in\mathcal{I}} \exp\left[\delta_h + \lambda^i_h\right]\right]$, to become:

$$\text{Prob}\left[\max_{i\in\mathcal{I}} : \quad \delta_h + \lambda^i_h + \varepsilon^i_h \leq \nu\right] = \exp\left[-e^{\widetilde{\nu}-\nu}\right].$$

It follows that the expected value of the maximum valuation of unit h is given by $\widetilde{\nu}$:

$$\mathcal{E}\left\{\widetilde{V}^{\mathcal{I}}_h(\mathbf{X}_h;\mathbf{Z})\right\} = \ell n\left(\sum_{i\in\mathcal{I}} \exp\left[\delta_h + \lambda^i_h\right]\right) = \delta_h + \ell n \int \exp\left[\lambda^i_h\right] \cdot dF_{\mathbf{Z}}, \qquad (8)$$

[4] See Anderson, de Palma and Thisse (1992). While it is noteworthy that under the assumption of the multinomial logit model, the expected maximum utilities, conditional on choice, are all equal to the overall expected maximum utility [Anas and Feng (1988)], this does not invalidate our approach. The multinomial logit model is invoked as a matter of analytical convenience here.

where $F_{\mathbf{Z}}$ denote the cumulative distribution function of the vector of characteristics of all individuals $(\mathbf{Z}_i)_{\mathcal{I}}$.

A number of remarks are in order. First, note that the above derivation of the hedonic price index is consistent with the standard definition of a hedonic index as the outer envelope of individual expenditure functions, where they are parameterized by income. Indeed, it does not depend on individuals' incomes. Second, note the symmetry between the expected value of maximum utility attained by a particular individual i, and the maximum valuation of unit h generated by the market, given by (7) and (8), respectively. Both involve averaging over characteristics, the former averaging over characteristics of units, $\mathbf{X}_h$, and the latter over characteristics of individuals, $\mathbf{Z}_i$.

To see the importance of selection correction bias that may underlie hedonic calculations, we compare the expected maximum valuation of unit h with the average valuation of unit h by the entire population of individuals. That is, compare (8) with the average value of (4) over the entire population of individuals:

$$\widetilde{\mathcal{V}}_h = \delta_h + \int \lambda_h^i \cdot \mathrm{Prob}\,_h^i \cdot dF_{\mathbf{Z}_i}. \tag{9}$$

Therefore, it follows that the average valuation understates the actual valuation, as implied by the choice model and given by (8). It is the latter that would be the outcome of the bidding process and should be the basis for the hedonic function that is associated with models of housing decisions involving choices over individual units.

Assuming alternative use of resources for supplying dwelling unit h into the market-allows us to fix the value, at equilibrium, of the left hand side of (8) above. Then, solving it in terms of p_h yields a hedonic price index for dwelling units with associated characteristics. Again, note that dwelling unit characteristics are interacted with the distributions of individual characteristics in arriving at the hedonic price function. Consider, for simplicity, that $\alpha_{jp} = 0, j \neq 0$. Then, (8) may solved to yield a hedonic price index:

$$\mathcal{E}\left\{ \widetilde{V}_h^{\mathcal{I}}(\mathbf{X}_h; \mathbf{Z}) \right\} = \ell n \left(\sum_{i \in \mathcal{I}} \exp\left[\delta_h + \lambda_h^i\right] \right) = \frac{1}{\alpha_{0_p}} \left[\alpha_{0X}\mathbf{X}_h + \xi_h + \ell n \int \exp\left[\lambda_h^i\right] \cdot dF_{\mathbf{Z}_i} \right]. \tag{10}$$

Therefore, the bias that excluding the individualized component of utility would cause on unit valuation, as in (6), is given by the last term in the above expression for the hedonic price index.

2.3 Overview of empirical findings by Bayer *et al.*

Bayer, McMillan and Rueben (2005) use the BLP approach, as explained above, by specifying utility from occupying a particular dwelling unit as a function of its price, of its characteristics and of characteristics of its neighborhood. The corresponding coefficients are individualized by being defined as explicit functions of individual characteristics. The model is applied to data at the level of the Census block from the San Francisco Bay area. Specifically, these authors use data from six counties in the San Francisco Bay area, which encompass about 650,000 individuals in 244,000 households who reside within 1,100 Census tracts that contain almost 39,000 Census blocks. The household data are obtained from the 1990 long Census forms and are geocoded down to the Census block level, a group of about 100 dwelling units. A vast amount of information that is available in the long Census forms and unavailable from other Census-based data sets, including incomes from a variety of sources as well as sociodemographic characteristics, are used in the estimations. These authors have complemented the restricted-access Census data with data from the complete set of housing transactions in the San Francisco Bay Area between 1992 and 1996. These data are based on county-level public records, and contain detailed information about every housing unit sold during that period, including the exact transaction price and the exact street address. The authors were able to add some additional characteristics by means of data collected in accordance with the Home Mortgage Disclosure Act (HMDA). These additional data allow them to investigate the robustness of their findings.

These authors embed the boundary discontinuity design (BDD) [Black (1999); Lee and Lemieux (2010)] in their estimation model by including school attendance zone boundary fixed effects in hedonic price regressions to control for the correlation of school quality and unobserved neighborhood quality. They extend Black's approach to deal with the systematic correlation of neighborhood sociodemographic characteristics and unobserved neighborhood quality, and to help identify the full distribution of household preferences for schools and neighbors. They provide new estimates of household preferences for schools and neighbors. Specifically, their key insight is that to the extent that one can control for differences in school quality on opposite sides of the a boundary, a boundary discontinuity design provides a plausible way to estimate the value that households place on the characteristics of their immediate neighbors. Using hedonic price regressions they show that the inclusion of boundary fixed-effects reduces the magnitudes of the coefficients on the income and education of one's neighbors by 25% and 60%, respectively. Thus higher-income and better-educated households seem to select into neighborhoods with better amenities. They also find that the coefficient on the fraction of black neighbors declines to zero, in sharp contrast to negative correlation of housing prices with the fraction of black neighbors typically observed and reported systematically in the previous literature. The authors argue that this is due to correlation of race and the unobserved neighborhood quality captured by the boundary fixed effect.

Generally, they find that hedonic price regression coefficients are generally very close to mean preferences for housing and neighborhood attributes that vary more or less continuously throughout the metropolitan area, including school quality and neighborhood income and education. In contrast, they find that estimated mean preferences for black neighbors differ markedly from hedonic estimates and are significantly negative. This underscores the importance of their method, since for blacks, who make up less than 10% of the population, mean and marginal households are far apart. Their analysis implies that, conditional on neighborhood income, households prefer to self-segregate on the basis of both race and education.

Such detail allows the authors to develop, *inter alia*, innovative new instruments for housing prices, in order to explore the economic effects of access to such neighborhood amenities as good schools and less crime, to examine in greater depth patterns in residential segregation [Bayer, McMillan and Reuben (2004a; 2004b; 2009)], and to estimate the full equilibrium effects of preferences for educational quality. That is, these effects include the direct valuation of educational quality as well as the indirect one via the fact that preferences for peers and neighbors who in turn (because of sorting) themselves value education more than the average individual. Bayer, Ferreira and McMillan (2007b) use general equilibrium simulations based on the estimates reported in Bayer, Ferreira and McMillan (2007a) to explore the size of social multiplier effects associated with increases in school quality as households re-sort. Such a social multiplier aspect of the preferences of neighborhood amenities and for school quality, in this particular case, is akin to the role of endogenous social effects more generally and has been overlooked by the literature.

2.4 Discrete location problems with endogenous and contextual effects

Bayer and Timmins (2005; 2007) emphasize problems associated with estimating models of location decisions in the presence of local spillovers measured by the number (and characteristics) of other agents (individuals or firms) who also choose the same or nearby locations at equilibrium. These are instances of endogenous social effects. As a number of authors have emphasized, using observations on patterns of locational decisions make it hard for a researcher to distinguish between the role played in the behavioral model by local spillovers, as opposed to underlying natural advantages, which may also be present and affect decisions. As they put it, in any model of sorting across locations, aggregate location decisions can be entirely accounted for by a vector of location-specific fixed-effects, which intermingle the influence of both natural advantages and local spillovers.

In order to disentangle the impact of unobserved local factors, Bayer and Timmins explore a source of potential instruments, which is similar to those used in the Bayer *et al.* work on housing decisions, which was discussed earlier. That is, they use functions of characteristics of other locations, which individuals could have chosen but did not, which are correlated with the fractions and characteristics of individuals who do choose particular locations but are uncorrelated with the unobserved fixed-attributes of those locations.

This particular location problem underscores the power of these types of instruments. However, it also serves to demonstrate the subtlety of the housing choice problem in the presence of neighborhood effects. As we demonstrate in Section 2.5, one of the difficulties of the housing decision is the fact that equilibrium-housing prices themselves are functions of contextual effects and endogenous variable in their own right.

This challenge is underscored by the similarity between the models involved. In particular, let us now index locations ℓ instead of dwelling units, as before, and rewrite (1) for the utility individual i derives from location ℓ as

$$V_\ell^i = \alpha_X^i X_\ell + \alpha_\varphi^i \phi_\ell + \xi_\ell + \varepsilon_\ell^i, \tag{11}$$

where ϕ_ℓ denotes the share of agents who choose location ℓ, and α_ϕ^i its respective coefficient, and the remainder of the notation is adapted in the obvious way from Section 2.1. As Bayer and Timmins (2007) indicate, inclusion of the location attribute ϕ_ℓ is meaningful when all other specific factors that may generate local spillovers that make a location attractive may not easily characterized structurally. The probability that individual i may choose location ℓ may be written as:

$$\text{Prob}_\ell^i = \mathcal{G}_{i\ell}(\mathbf{Z}^i, \mathbf{X}, \Phi, \Xi), \tag{12}$$

where $\mathbf{Z}^i, \mathbf{X}, \bar{\phi}, \bar{\xi}$ are the matrix/vector counterparts of the respective variables. By aggregating over all individuals we have $\phi_\ell = \sum_{i \in \mathcal{I}} \int \mathcal{G}_{i\ell}(\mathbf{Z}^i, \mathbf{X}, \bar{\varphi}, \bar{\xi}) df(\mathbf{Z})$, from which we may write for all locations, $\ell \in \mathcal{L}$:

$$\Phi = \mathcal{G}(f(\mathbf{Z}, \mathbf{X}, \Phi, \Xi). \tag{13}$$

This system of equations maps $[0,1]^L$ into itself, where $L = |\mathcal{L}| = |\cup_\ell \mathcal{L}_i|$, is the size of the union of all individuals' discrete choice sets, $\mathcal{L}_i$.

We note that in moving from the individual decision, denoted by (12), to the aggregate (13), we have assumed in effect that all agents' location decisions form a static simultaneous move game, whose Nash equilibria are represented by the fixed-points of (13). The game assumes that agents have full information about each other's unobserved components in the preference structure, $\bar{\varepsilon}^i = (\ldots, \varepsilon_h^i, \ldots)$. Alternatively, we may assume that agents form expectations about others' preference parameters. Under the assumption of a continuum of agents with different values of unobserved parameters these terms may be integrated out, thus avoiding integer problems.

Similarly to the housing estimation problem discussed above, Bayer and Timmins propose a two-stage estimation. In the first stage, the vector of location-specific constants, $\mathcal{D} = (\ldots, \delta_\ell, \ldots)$, is estimated, where

$$\delta_\ell = \alpha_{0X} X_\ell + \alpha_{0\phi} \phi_\ell + \xi_\ell, \tag{14}$$

$(\alpha_{00X}, \ldots, \alpha_{0rX}, \ldots)$ denotes an array and $\alpha_{0\phi}$, a scalar, and a stochastic structure for the ε_ℓ^i's, the shocks in (11) is assumed. These shocks are assumed to be IID across individuals according to a distribution $F(0, \Sigma)$, where the covariance is defined over locations. In the second stage, the estimated vector $\hat{\mathcal{D}}$ is regressed against attributes of locations X_ℓ and ϕ_ℓ. However, the logic of the location decisions implies that ϕ_ℓ and ξ_ℓ are correlated and therefore, conditions that would allow us to recover parameters of interest by regressing the estimated fixed-effects are not satisfied.

Bayer and Timmins propose as instruments the exogenous attributes of location ℓ and of other locations which the agents could have chosen but did not do so. They motivate this choice by what they view as the optimal instrument, the expected share of individuals who locate at each site, conditional on the full set of exogenous location and individual characteristics, $\mathcal{E}(\phi_\ell|\mathbf{X}, \mathbf{Z})$. Specifically, they propose as an instrument the predicted share of each alternative, to be evaluated at the first stage estimates of the parameter vector while the unobserved location specific components in Equation (14) $\Xi = (\xi_1, \ldots, \xi_\ell, \ldots, \xi_L)$, and the coefficients of the local spillover terms, $\alpha_{0\phi}$, are set equal to zero and only the observed exogenous choice and individual characteristics are being considered. This neglects the role of local spillovers but provides a measure of the way that the full landscape, which is based on exogenous features that are observable, of possible choices enters the problem. That is, the estimation would be based on the assumption that each individual behaves optimally, given the collection of choices made by other individuals. We note below that the estimation process in Ioannides and Zabel (2008) is quite similar in spirit and execution.

In view of the properties of the individual choice model, Equation (13) may admit a multiplicity of equilibria, as Brock and Durlauf (2001, 2007) have emphasized in similar settings involving social interactions. Bayer and Timmins (2005), Propositions 2 and 3, propose conditions that ensure uniqueness. When such conditions are implausible, one would need to interpret cautiously the associated identification issues [Bisin *et al.* (2009)]. These authors evaluate the efficiency and computational feasibility of different approaches to solving the curse of dimensionality implied by the equilibrium multiplicity. While multiplicities are troublesome in certain settings, they are important in other ones, and their empirical significance has been overlooked in the literature. In particular, Brock and Durlauf (2007) develop partial identification results for binary choice models under what they argue are weak assumptions about the distribution of the unobserved group effects. They exploit a fundamental difference between endogenous effects and unobserved group effects: only endogenous effects can produce multiple equilibria. Hence, if evidence of multiple equilibria can be generated, it would represent evidence of endogenous social interactions. Brock and Durlauf consider "pattern reversals" in group level outcomes. A pattern reversal occurs when the rank order of average outcomes between two groups is the reverse of what one would predict given the observed individual and contextual effects for the groups. Brock and Durlauf demonstrate that under various shape restrictions on the probability density of the unobservables, pattern reversals can only occur because of multiple equilibria and hence endogenous social interactions. For example, if the distribution of unobservables shifts monotonically in observables, then pattern reversals cannot occur without social interactions. A second approach they propose involves the bimodality of linear combinations of contextual effects. Here Brock and Durlauf show that, conditioning on a given average outcome, the cross-section distribution of certain linear combinations of contextual variables must be unimodal.

2.5 Endogenous contextual effects and hedonic prices: the Nesheim model

Here we take up the solution to a model of sorting, originally due to Nesheim (2002), that allows for contextual effects, own income effects, and endogenous choice of neighborhood. Consequently, one can use it to study properties of entire neighborhoods to which individuals self-select, that is sort. It is important to account for the roles that contextual effects, measured at different scales, play in individuals' decisions. The theory we turn to next allows for a better understanding of the fundamental economic and social forces that hold together groups of individuals.

Following Nesheim (2002), I assume that individuals choose their residential location from among a continuum of locations indexed by ℓ. Individuals consume a unit of housing each and have preferences over nonnonhousing consumption and their own child's expected schooling. Locations differ in terms of an attribute, average neighborhood schooling among adults, that individuals value as an input to their children's education. Each individual (household) allocates her income over nonnonhousing consumption, $I-R(\ell)$, where I and $R(\ell)$, are, respectively, income and unit housing rent at location $\ell, \ell \in R_+$, and housing, $R(\ell)$.

I describe households in terms of a vector of attributes, $\mathbf{z} = (z_1, z_2, z_3, z_4, z_5)$, whose components are defined respectively as: log of parental schooling, $s_0 = e^{z_1}$, log of own parent's income, $I = e^{z_2}$, log of the child's ability in school, $a = e^{z_3}$, log of a preference parameter, $\beta = e^{z_4}$, that weights a child's schooling outcome in the utility function, and a random shock to schooling outcome, z_5, which will be assumed to be uncorrelated with all other components:

$$\mathbf{z} = (\ell n s_0, \ell n I, \ell n a, \ell n \beta, z_5).$$

Households maximize utility by choosing location, ℓ. Utility is additively separable in nonhousing consumption, $I-R(\ell)$, and in expected schooling for the child, conditional on parental characteristics and location, $\mathcal{E}(s_1|\ell)$:

$$\max_{\ell} : \left\{ \frac{1}{\gamma} \left[1 - e^{-\gamma(I-R(\ell))} \right] + e^{z_4} \mathcal{E}(s_1|\ell) \right\}, \gamma \geq 0. \tag{15}$$

A child's schooling is produced in location ℓ as described by an educational production function as a function of: average schooling in ℓ, $S(\ell)$; own parent's schooling, e^{z_1}; the child's ability, e^{z_3}; and the random shock e^{z_5}. That is:

$$s_1 = (S(\ell))^{\eta_1} e^{\eta_2 z_1 + z_3 + z_5}, \tag{16}$$

where η_1 and η_2 are positive parameters. The random shock z_5 is the only quantity that is unobservable by the household when it chooses location and is assumed to be independent of location. Average schooling at each location, $S(\ell)$, and the housing rent, $R(\ell)$, are both endogenous. They are determined consistently with equilibrium sorting of households across locations.

The sorting equilibrium is defined in terms of: one, a mapping $F(\mathbf{z})$ that assigns household types $\mathbf{z}$ to locations, $\mathbf{z} \in Z$, $F(z)$: $Z \to R_+$; and two, a housing rent function $R(\ell)$, that equilibrates the housing market in each location, $\ell \in R_+$. $S(\ell)$, average schooling of parents who choose to locate in neighborhood $\ell, S(\ell)$, is defined as $S(\ell) = \mathcal{E}[e^{z_1} | z \in F^{-1}(\ell)]$. This is the location-specific input into the education production function (16).

Provided that monotonicity holds at equilibrium, it is convenient to represent location in the utility maximization problem in terms of the average schooling of parents in each neighborhood, $S(\ell)$. Consequently, the housing rent function is equivalently defined as a function of the contextual effect, S, $p(S)$, that is instead of rent at location ℓ.

Nesheim (2002) shows that when $\gamma \neq 0$ the differential equation that the hedonic price function must satisfy, does not allow an analytical solution and may only be solved numerically. Nesheim establishes several general results, such as existence of a unique equilibrium rent function, provided that groups of individuals with higher average willingness to pay for education are willing to pay more for high quality locations with higher education [*ibid.*, Theorem 5.1]. However, it is possible that the average education of people who choose a location with a given quality index may decline with the value of the quality index. In that case, Theorem 5.2, *ibid.*, ensures the existence of a nontrivial equilibrium rent function with a discontinuous slope. Accordingly, there will be concentrations of individuals of finite mass at the points of discontinuity. Nesheim shows that this more general model is econometrically fully identified.

2.6 A generalization of the Nesheim sorting model

It turns out that by modifying a single assumption in Nesheim (2002), namely that component z_2 of the vector of individual characteristics $\mathbf{z}$ is the *level* of parental income, *not* its *log*, allows an analytical solution for $p(S)$. [Ioannides (2008)] The first order condition for maximizing utility (15) with respect to S is:

$$-e^{-\gamma[z_z - p(s)]} p_s(S) + A_0 \eta_1 S^{\eta_1 - 1} e^{\eta_2 z_1 + z_3 + z_4} = 0. \tag{17}$$

Rearranging and taking logs yields

$$\eta_2 z_1 + \gamma z_2 + z_3 + z_4 = \Theta(S), \tag{18}$$

where the marginal quality index is defined as:

$$\Theta(S) \equiv \gamma p(S) + \ell n[p_s(S)] - \ell n(A_0\eta_1) + (1-\eta_1)\ell n S, \tag{19}$$

with $A_0 = \mathcal{E}(e^{z_5})$. This is the marginal effect of neighborhood quality on utility, given its price. The l.h.s. of (18) expresses the marginal willingness to pay for neighborhood quality.

Let $\varphi\ \varphi^{\mathbf{T}} = (\eta_2, \gamma, 1, 1, 0)$ denote a vector of parameters. We complete the description of the problem by defining average neighborhood schooling, that is, the average schooling of adults who choose location S:

$$S = \mathcal{E}\{e^{z_1} | \varphi\mathbf{z} = \Theta(S)\}. \tag{20}$$

Those who choose a particular neighborhood are defined in terms of their characteristics, as implied by (18). If $\mathbf{z} \sim \mathcal{N}(\mu, \Sigma)$, then the log education of parents who choose location S is by the properties of the multivariate normal distribution normally distributed: $(z_1|\varphi^{\mathrm{T}} z = \Theta(S)) \sim \mathcal{N}(\widetilde{\mu}_1, \widetilde{\sigma}_1^2)$, with mean and variance given by

$$\widetilde{\mu}_1 = \mu_1 + \frac{\varphi^{\mathrm{T}}\Sigma e_1}{\varphi^{\mathrm{T}}\Sigma\varphi}\left[\Theta(S) - \varphi^{\mathrm{T}}\mu\right], \quad \widetilde{\sigma}_1^2 = \sigma_{11}^2\left(1-\widetilde{\rho}_1^2\right);$$

where: $\widetilde{\rho}_1^2 = \frac{(\varphi^{\mathrm{T}}\Sigma e_1)^2}{\sigma_{11}(\varphi^{\mathrm{T}}\Sigma\varphi)}$, $e_1^{\mathrm{T}} \equiv (1,0,0,0,0)$, denotes the square of the correlation coefficient between log education of parents and willingness to pay for neighborhood quality. Clearly, due to sorting, the variance of the log education of parents, conditional on the willingness to pay for neighborhood quality is less than in the entire population.

Using definition (20) with the first-order condition for utility maximization (17), integrating, and using the initial condition $p(0) = 0$, yields the equilibrium price function:

$$p(S) = \frac{L_1}{\gamma}\ell n\left[1 + \frac{\gamma}{L_1} L_0^{-\frac{1}{L_1}}\left(\eta_1 + \frac{1}{L_1}\right)^{-1} S^{\eta_1 + \frac{1}{L_1}}\right], \tag{21}$$

where the auxiliary variables L_0, L_1, as functions of preference parameters and of the parameters of the joint distribution of vector $\mathbf{z}$, are defined as follows:

$$L_0 \equiv e^{\mu_1 - L_1[\ell n(\eta_1 A_0) + \varphi^{\mathrm{T}}\mu)] + \frac{1}{2}\widetilde{\sigma}_{11}^2}, \quad L_1 \equiv \frac{\varphi^{\mathrm{T}}\Sigma e_1}{\varphi^{\mathrm{T}}\Sigma\varphi}.$$

Note that L_1 is the regression coefficient from the regression of log-education of parents on the willingness to pay for neighborhood quality [*c.f.*, Nesheim, *op. cit.* p. 33].

2.7 Properties of the hedonic price function

The equilibrium price function is increasing in S. Analysis of its second derivative shows that the equilibrium price function is convex for low values of S and up to a threshold point, and concave thereafter, provided that $\eta_1 + \frac{1}{L_1} > 1, \gamma > 0$. Thus, the equilibrium rent function exhibits a key property of social interactions models: the equilibrium rent is a sigmoid function of average neighborhood quality.

In view of the definition of L_1 above, the regression coefficient from the regression of log-education of parents on the willingness to pay for neighborhood quality, this condition involves restrictions in terms of parameters and of the stochastic structure:

$$\eta_1 + \frac{\phi^{\mathrm{T}}\Sigma\phi}{\eta_2\sigma_{11} + \gamma\sigma_{21} + \sigma_{31} + \sigma_{41}} > 1.$$

As expected from hedonic price theory, an individual's marginal valuation of neighborhood quality, the hedonic demand equation (18), depends on the *individual's* income, but the equilibrium rent function (21) depends, via the auxiliary functions L_0, L_1, on the *statistics* of the income distribution and its joint distribution with the other characteristics of interest.

Average neighborhood quality; that is schooling, chosen by an individual with characteristics (z_1, z_2, z_3, z_4) follows by using solution (21) in (17). That is:

$$\left[1 + \frac{\gamma}{L_1} L_0^{-\frac{1}{L_1}} \left(\eta_1 + \frac{1}{L_1}\right)^{-1} S^{\eta_1 + \frac{1}{L_1}}\right]^{L_1 - 1} S^{\frac{1}{L_1}} = \widetilde{A}_0 e^{\eta_2 z_1 + \gamma z_2 + z_3 + z_4},$$

where $\widetilde{A}_0 \equiv A_0 \eta_1 L_0^{\frac{1}{L_1}}$. The choice of neighborhood for an individual with characteristics (z_1, z_2, z_3, z_4) is implicitly determined by the roots of the above equation. Given a solution for S, $S = \mathcal{S}(z_1, z_2, z_3, z_4)$. Then a child's education follows from (16), $\ell n s_1 = \eta_2 z_1 + z_3 + z_5 + \eta_1 \mathcal{S}(z_1, z_2, z_3, z_4)$. The dependence of a child's education on z_2, parental income, is brought in via $\mathcal{S}$.

The magnitude of L_1 does affect the sensitivity of S with respect to the willingness to pay for neighborhood quality. If $L_1 < 1$ and $\eta_1 + \frac{1}{L_1} < 1$, then there will, in general and subject to feasibility conditions, exist two solutions. For only one of them, the smaller in magnitude, it would be the case that neighborhood education increases with willingness to pay for it and with income, in particular. This is also the case if $\eta_1 + \frac{1}{L_1} > 1$, and provided the value is not too large. Both solutions are in principle acceptable, but they have different properties. For example, the larger of the two solutions implies that individuals with more own schooling choose neighborhoods with lower average schooling, which in turn implies lower schooling for their children. In other words, depending upon parameter values, the model allows for schooling to be either a normal or an inferior good.

Another example of the fruitfulness of this approach is to consider the distribution of income, conditional on location choice S. The income of people who choose S is normally distributed:

$$(z_2|\varphi^{\mathrm{T}}z = \Theta(S)) \sim N(\widetilde{\mu}_2, \widetilde{\sigma}_{22}),$$

with mean and variance given by:

$$\widetilde{\mu}_2 = \mu_2 + \frac{\varphi^{\mathrm{T}}\Sigma e_2}{\varphi^{\mathrm{T}}\Sigma\varphi}\left[\Theta(S) - \varphi^{\mathrm{T}}\mu\right], \;\; \widetilde{\sigma}_2 = \sigma_{22}\left(1 - \widetilde{\rho}_2^2\right),$$

where $\widetilde{\rho}_2^2 = \frac{(\varphi^{\mathrm{T}}\Sigma e_2)^2}{\sigma_{22}(\varphi^{\mathrm{T}}\Sigma\varphi)}$, denotes the square of the correlation coefficient between the income of parents and the willingness to pay for neighborhood quality, and solution (21) is used in the expression for $\Theta(S)$, in (19). In this case, function $\Theta(S)$ becomes:

$$\Theta(S) \equiv -\ell n(\widetilde{A}_0) + \frac{1}{L_1}\ell nS + (L_1 - 1)\ell n\left[1 + \frac{\gamma}{L_1}L_0^{-\frac{1}{L_1}}\left(\eta_1 + \frac{1}{L_1}\right)^{-1}S^{\eta_1 + \frac{1}{L_1}}\right].$$

The conditional distribution of education of parents who choose S has a lognormal distribution, $(z_1|\varphi^{\mathrm{T}}z = \Theta(S)) \sim N(\widetilde{\mu}_1, \widetilde{\sigma}_{11})$, where

$$\widetilde{\mu}_1 = \mu_1 + \frac{\varphi^{\mathrm{T}}\Sigma e_1}{\varphi^{\mathrm{T}}\Sigma\varphi}\left[\Theta(S) - \varphi^{\mathrm{T}}\mu\right], \;\; \widetilde{\sigma}_1 = \sigma_{11}\left(1 - \widetilde{\rho}_1^2\right),$$

and $\widetilde{\rho}_1^2 = \frac{(\varphi^{\mathrm{T}}\Sigma e_1)^2}{\sigma_{11}(\varphi^{\mathrm{T}}\Sigma\varphi)}$, denotes the square of the correlation coefficient between log education of parents and willingness to pay for neighborhood quality.

2.8 The special case of no income effects

Setting $\gamma = 0$ in (21) yields the hedonic price of housing obtained by Nesheim:

$$p(S) = L_0^{-\frac{1}{L_1}}\left(\eta_1 + \frac{1}{L_1}\right)^{-1}S^{\eta_1 + \frac{1}{L_1}}. \tag{22}$$

This solution is also implied by (21), for small values of S.[5] The elasticity of the equilibrium rent function with respect to neighborhood quality S, in the absence of income effects, $\eta_1 + \frac{1}{L_1}$, is intuitive. The larger is η_1, the larger price differentials must be to segregate individuals into their preferred locations. The larger is L_1, the regression coefficient of log-education of parents on the willingness to pay for neighborhood quality, the less parents are willing to pay directly for neighborhood quality, since they contribute indirectly through their own education and therefore the smaller the price

[5] For large values of S, the second term within the brackets in the RHS of (21) dominates the first, thus admitting an approximation: $p(S) \approx \frac{L_1}{\gamma}\;\ell n\left[\frac{\gamma}{L_1}L_0^{-\frac{1}{L_1}}\left(\eta_1 + \frac{1}{L_1}\right)^{-1}\right] + \frac{1+\eta_1 L_1}{\gamma}\ell nS.$

differentials are required to be to maintain segregation of households at equilibrium. The equilibrium price function (22) is convex (concave), if $\eta_1 + \frac{1}{L_1} > (<) 1$. However, for the solution to be meaningful, $\eta_1 + \frac{1}{L_1} > 0$.

Such endogenous distributions of interest as those of education and of income of parents who choose neighborhood education S readily follow. First, neighborhood quality as a function of S becomes: $\Theta(S) \equiv -\frac{1}{L_1} \ell n L_0 - \ell n(A_0\eta_1) + \frac{1}{L_1} \ell n S$. So, we have: $S = L_0^{-\frac{1}{L_1}} (A_0\eta_1)^{L_1} e^{L_1(\eta_2 z_1 + z_3 + z_4)}$. The log of average neighborhood schooling *chosen* by individual with characteristics $\mathbf{z}$ is linear in $L_1(\eta_2\, z_1 + z_3 + z_4)$, where the effect of individual characteristics is moderated by L_1. The larger is L_1, the regression of log-education of parents on the willingness to pay for neighborhood quality, the more neighborhood quality parents are willing to purchase directly, since this makes their own education more effective in ensuring greater segregation of households at equilibrium.

Also interesting is the relationship between the child's education and household characteristics. That is, from (16), by using the solution for S yields:

$$\ell n s_1 = a_0 + \eta_2(1 + \eta_1 L_1) z_1 + (1 + \eta_1 L_1) z_3 + \eta_1 L_1 z_4 + z_5, \tag{23}$$

where a_0 is a function of parameters. The intergenerational evolution of schooling is determined by $\eta_2(1 + \eta_1\, L_1)$, the coefficient of the log of parents' schooling in (23).

It is now clear how restrictive the typical hedonic price literature approach is, where estimates are sought for *arbitrary* functional forms for the relationship between the (marginal) valuation for neighborhood amenities as a function of observables and unobservables. In contrast, Nesheim's theory exactly determines the equilibrium rent, as a function of contextual characteristics, and this in turn characterizes sorting. As Nesheim (2002) shows, this particular model is not entirely identified econometrically, although groups of parameters may be identified by means of data on observable educational outcomes as a function of parental education, neighborhood school quality, and income.

If individuals value the characteristics of their neighbors, then as an outcome of their choice of location their neighbors' characteristics are correlated with their own. Similarly, if certain individual outcomes that households care about, like educational attainment of children, depend on characteristics of their neighbors, then unobservable characteristics of individuals are likely to be correlated with neighborhood characteristics. If children's ability increases the productivity of neighborhood quality in producing education, then people who know they have high ability children will move to high quality neighborhoods. This leads to sorting: higher quality neighborhoods will have higher unobserved ability.

Such neighborhood effects on schooling have been posited as stylized facts and addressed by typically atheoretical papers, such as Brooks–Gunn *et al.* (1993), who emphasize the importance of sorting bias, and Kremer (1997), who estimates schooling as a function of parental schooling and neighborhood schooling and assesses the implied role of sorting in inequality. Ioannides (2003) shows empirically that nonlinearities in the general relationship, whose special case is estimated by Kremer (1997), may alter his key results. Graham (2008) seeks to identify the causal effects of parental schooling and neighborhood schooling, while recognizing that parents choose the neighborhoods where they bring up their children. His model of intertemporal optimization yields an Euler-like equation, which links the hedonic price of housing to parental schooling, neighborhood schooling and total ability and is the counterpart of (19). Its integration yields a hedonic price as an isoelastic function of neighborhood schooling, an exact counterpart of (22), and is thus less general than my generalization of Nesheim's example (21). In Section 5.2, we discuss broadly related results that are obtained by means of an Epple-type model [Calabrese *et al.* (2006)].

2.8.1 Housing price as an empirical control for neighborhood quality

Nesheim (2002) reports estimation results with data from the US National Educational Longitudinal Survey for 1998. He demonstrates the key difficulty in estimating the model, the dependence at equilibrium between marginal price and neighborhood quality. However, use of different data sets, which can allow controlling for additional determinants of housing price, can facilitate identification. For example, the confidential version of the Panel Study of Income Dynamics links individual and parental characteristics to the characteristics of the neighborhoods where respondents grew up and thus allows for additional information on housing price determinants to be brought to bear on the estimation. Alternatively, regression discontinuity designs, like that employed by Black (1999), can also be helpful, precisely because they allow for exogenous instruments.

Conceptually related is Bayer and Ross (2009) who use structural features of sorting mechanisms, as analyzed by the Epple *et al.* class of models and discussed in detail in Section 5.2. By choosing where to live, individuals may be motivated not only about attributes of houses and general neighborhood amenities, but also specifically by how their labor market outcomes and therefore income prospects would be affected. This choice makes labor market outcomes depend on individual and neighborhood observable and unobservable variables see also Bayer, Ross and Topa (2008). These authors exploit the monotonic relationship between neighborhood housing prices and neighborhood quality as an empirical control function for the neighborhood unobservable in the determination of labor market outcomes. This device in effect transforms the problem to a model with one unobservable so that traditional instrumental variables solutions may be applied. Mechanically speaking, this approach is quite similar to controlling for self-selection bias associated with neighborhood choice, as in the Nesheim model and its generalization in Section 2.6. Bayer and Ross instrument for each individual's observed neighborhood attributes with the average neighborhood attributes of a set of observationally identical individual's which they obtain by conditioning with respect to a variety of controls available in the data.

They estimate a model of labor market outcomes as a function of neighborhood effects using confidential micro data from the 1990 Decennial Census for the Boston MSA. The outcomes used include: labor force participation, weeks worked and weekly hours in the previous year. The results imply that the direct effects of geographic proximity to jobs, neighborhood poverty rates, and average neighborhood education are substantially larger than the conditional correlations identified using OLS, although the net effect of neighborhood quality on labor market outcomes remains small. This suggests that individuals with a lower likelihood of obtaining employment have sorted into locations with superior labor market opportunities potentially to compensate for their poor unobservables.

Specifically, the authors find that neighborhoods have large and complex effects on labor market outcomes. Employment access, low levels of poverty, a low fraction of college graduates, and high levels of unobserved neighborhood attributes are all associated with higher levels of labor force participation, greater number of weeks worked in a year, and with the exception of poverty greater average number of hours worked per week. For example, a one standard deviation increase in employment access leads to approximately a four percentage point increase in labor force participation in the sub-sample of individuals who have never attended college. The positive impact of good job access on the intensity of labor force participation, as captured by weeks per year and hours per week, does not change as the human capital level of the sample falls. The effects regarding gender are even more striking: all findings decline in magnitude and many become statistically insignificant as married women and eventually all women are deleted from the sample. Overall, the results indicate that neighborhood effects are most important for individuals with weak attachment to the labor market, especially married women. Still, while the effect of individual variables appears large, the net effect of neighborhood quality is small. Neighborhoods with low poverty rates and other attributes that positively impact labor market outcomes appear correlated with the percent of college graduates in equilibrium. These competing effects lead to small and sometimes negative relationships between overall neighborhood quality and various labor market outcomes. This is consistent with findings in research based on data from the Moving to Opportunity program that improvements in neighborhoods quality had little or no impact on earnings.

3. ENDOGENOUS NEIGHBORHOOD AND CONTEXTUAL EFFECTS IN HOUSING MARKETS

In the context of housing and other decisions, close physical and "mental" proximity is likely to be associated with individuals' caring about the actual or expected behavior of their neighbors. These are individuals' decisions and thus by definition they generate endogenous social effects. Because of sorting into neighborhoods, neighbors' characteristics are likely to be correlated. Results by Kiel and Zabel (2009) confirm this intuition using data from the national sample of the American Housing Survey (NAHS).[6] The correlation coefficient between nonwhite household head and the percentage of nonwhite heads in the immediate neighborhood and in the census tract are 0.749 and 0.677, respectively, and for the percentage of nonwhite heads between immediate neighborhood and tract 0.885. Similarly, the corresponding estimates for share with completed high school are 0.331, 0.276, and 0.591, and for permanent income 0.648, 0.451, and 0.607, respectively.

In order to understand better such outcomes, we augment individuals' utility functions in the following fashion. Individual i is assumed to care about housing services indirectly produced by a vector of characteristics of her dwelling unit h, such as size, number of rooms, number of baths, age, etc., which are to be denoted by $\mathbf{x}_h$ and referred to as *dwelling attributes*, and by attributes $\mathbf{x}_{\ell(h)}$ of the neighborhood, which includes location within an urban area, typically identified via the census tract ℓ or the community (governmental jurisdiction) in which it is located, to be referred to as *neighborhood attributes*. These may include summary statistics for the $\mathbf{x}_h$'s of dwelling units belonging to neighborhood ℓ and, in addition, housing quality attributes of the surrounding neighborhood itself that is comprised of immediate neighbors, etc., to be denoted by k, as well as other characteristics of a tract, $\mathbf{g}_\ell$. In addition, individual i is assumed to care about observable and unobservable demographic characteristics $\mathbf{z}_{k(i)}$, among individual i's immediate neighbors in k, or in $\ell(i)$, the census tract where she resides. These are referred to as *contextual (or exogenous social) effects* in the remainder of the chapter. Individual i is also assumed to care about the vector of housing consumptions in her immediate neighborhood, $\mathbf{Y}_{k(i)}$. This feature will be referred to as *endogenous social effect*. A utility function $V_{\ell ki}$ is specified as:

$$V_{\ell ki} = U\left(c_i, \left(\mathbf{x}_h, \mathbf{x}_k, \mathbf{g}_\ell\right); \mathbf{z}_i; \mathbf{Y}_{k(i)}, \mathbf{z}_{k(i)}\right). \tag{24}$$

We assume that each of the vectors of dwelling and of neighborhood attributes may be partitioned into two sub-vectors, with the first corresponding to size (scale) related attributes and the second to potentially scale free ones (like ambience, reputation, etc.) that characterize neighborhood quality. We posit a scalar measure of housing services, defined as a function $\mathcal{Y}(\mathbf{X}_{stru(h)}, \mathbf{X}_{\ell(h)})$, the vectors of dwelling unit h and neighborhood attributes of its location, $\ell(h)$, respectively. It is appropriate to think

[6] The micro data from the American Housing Survey were geocoded by means of privileged access to confidential US Census data. The main data source used for this study is the neighborhood clusters subsample of the national sample of the American Housing Survey (NAHS). The NAHS is an unbalanced panel of more than 50,000 housing units that are interviewed every two years and contains detailed information on dwelling units and their occupants through time, including the current owner's evaluation of the unit's market value. This data set is also used by Ioannides and Zabel (2008) and is discussed further below.

of such a scalar measure of housing services as an *index*, and I return to such an interpretation further below.

Appealing to a theorem of Samuelson and Swamy (1974) on consistent aggregation [see also Sieg *et al*. (2002)] we posit that the size-related attributes of dwelling units enter preferences through a separable function for housing services $\mathcal{Y}(\cdot)$, that is homogeneous of degree one in all inputs.[7] We specify $\mathcal{Y}(\cdot)$ as follows:

$$Y_h = \mathcal{Y}(\mathbf{X}_{\ell stru(h)}, \mathbf{X}_{\ell nei(h)}) \equiv x_{stru(h)}^{1-\vartheta} x_{\ell(h)}^{\vartheta}, \quad 0 < \vartheta < 1, \tag{25}$$

where inputs $(x_{stru(h)}, x_{\ell(h)})$ are scalar, for simplicity, and ϑ a parameter. The corresponding sub-expenditure function for housing can be defined as the minimum cost necessary to obtain Y_h units of housing services once prices for the two types of inputs are given, $\left(P_{\ell,stru(h)}, P_{\ell,nei(h)}\right)$. That is:

$$E_Y = (1-\vartheta)^{-(1-\vartheta)} \vartheta^{-\vartheta} P_{stru(h)}^{1-\vartheta} P_{\ell(h)}^{\vartheta} \cdot Y_h. \tag{26}$$

A definition of a price index for housing services from a dwelling unit in ℓ readily follows:

$$P_\ell \equiv (1-\vartheta)^{-1+\vartheta} \vartheta^{-\vartheta} P_{\ell,stru}^{1-\vartheta} P_{\ell,nei}^{\vartheta}. \tag{27}$$

By using (25) in (26) and taking logs of both sides of (26) above yields a regression-like equation

$$\ell n E_Y = \ell n P_\ell + (1-\vartheta) \ell n x_{stru(h)} + \vartheta x_{\ell(h)}.$$

Suppose that for a sample of dwelling units in different locations (communities) housing expenditure E_Y and $x_{stru(h)}$ are observable but not $x_{\ell(h)}$. Sieg *et al*. (2002) argue that community-specific housing prices P_ℓ are identified (up to scale) by the tract- or community-specific fixed-effects in the regression model. More generally, $\mathbf{X}_{stru(h)}$ can be a vector of observable housing characteristics, instead of the scalar $x_{stru(h)}$ above, and all unobservable attributes are included in the error. This definition may be used to express the quantity of housing services (an inherently unobservable quantity)[8] as

[7] We note, however, that the linear homogeneity restriction of housing services (26) is arbitrary in the context of the Ekeland *et al*. (2004) hedonic theory and of Nesheim (2002), discussed above. Yet, it is suggested by the Samuelson and Swamy, *op. cit*., requirements for consistent aggregation that imply an invariant price index. It may be treated as an overidentifying restriction that can be tested by means of techniques similar to measures of rank violations that are used in comparing different types of indices that Epple, Sieg and their co-authors themselves employ.

[8] This feature of housing as a commodity is well known, of course, and numerous approaches have aimed at circumventing it. The first systematic one, however, is Epple, Gordon and Sieg (2009), which estimates housing production functions when both housing quantities and prices are treated as unobserved latent variables. The approach rests on duality theory and incorporates the value of housing per unit of land and the price of land in order to obtain an alternative representation of the indirect profit function as a function of those observables. The empirical demonstration of their approach uses data from recently built properties in Allegheny County, Pennsylvania. An estimated Cobb-Douglas production function has a land share of 0.144 and thus a share of 0.856 for nonland mobile factors.

a function of the observable housing expenditure E_Y and of the components of the price index, which may be estimated.[9]

3.1 An application: the Ioannides and Zabel model of neighborhood choice and housing demand as a joint decision

Ioannides and Zabel (2008) develop a model of housing structure demand with neighborhood effects and of neighborhood choice as a joint decision. The estimation exploits a household-level data set that has been augmented with contextual information at several levels ("scales") of aggregation. One is at the neighborhood level, consisting of about ten immediate neighbors and using data from the neighborhood clusters sub-sample of the American Housing Survey. Another level is the census tract to which these dwelling units belong.[10] A third one is the metropolitan area in which the respective census tracts lie.

3.1.1 The preference structure

We follow Ioannides and Zabel (2008) and assume household i chooses among dwelling units that belong to neighborhood cluster k, $k = 1, \ldots, K_s$, in tract $\ell, \ell = 1, \ldots, L$, in a given metropolitan area, and specify the utility function $\mathbf{\Omega}_{\ell ki}$, as a variation of (24) to be made up of two multiplicative components. The first component, $V_{\ell ki}$, is a *conditional* indirect utility function, that is specified below as a function of prices, income, and additional observable and unobservable characteristics of individuals residing in neighborhood cluster k and in the census tract ℓ in which k lies. The second, $\varepsilon_{\ell ki}$, is a random component of utility, drawn from a distribution to be specified below, that affects neighborhood choice and is assumed to be observable by the individual and unobservable by the econometrician.[11] We specify the conditional indirect utility function as being made up of a component reflecting tract-specific characteristics, $\mathbf{g}_\ell$, and of a component reflecting the value to household i from nonhousing consumption and consumption of housing services, $\omega_{\ell ki}, \exp[\zeta_i \mathbf{g}_\ell] \cdot \omega_{\ell ki}$. That is:

$$\mathbf{\Omega}_{\ell ki} = V_{\ell ki} \cdot \exp[\varepsilon_{\ell ki}] = \exp[\zeta_i \mathbf{g}_\ell] \cdot \omega_{\ell ki} \cdot \exp[\varepsilon_{\ell ki}]. \tag{28}$$

The term $\omega_{\ell ki}$ is defined as the maximum value of a direct utility function with respect to nonhousing consumption, c_i, and consumption of housing services, Y_i, subject to a budget constraint, $c_i + P_\ell \cdot Y_i = I_i$, where I_i denotes household income and P_ℓ the housing price: $\omega_{\ell ki} \equiv \exp\left[\frac{I_i^{1-\delta}-1}{1-\delta}\right] \cdot \exp\left[-\frac{B_{ki} P_\ell^{\mu+1}-1}{\mu+1}\right]$. That is:

[9] However, the above definition is appropriate when both components are freely variable. Once a component has been fixed, the appropriate expenditure function is the quasi-fixed one, and thus different from (26).

[10] These authors define neighborhood selection as discrete choice over census tracts. Public Census data provide information on the joint distribution of various variables within tracts, which is crucial for the estimation by Ioannides and Zabel and by Epple and co-authors, which we discuss in section 5.1 below.

[11] This model is influenced by features of Dubin and McFadden (1984) and of Epple, Romer and Sieg (2001).

$$\Omega_{\ell ki} \equiv \exp\left[\zeta_i \mathbf{g}_\ell\right] \cdot \exp\left[\frac{I_i^{1-\delta}-1}{1-\delta}\right] \cdot \exp\left[-\frac{B_{ki}(\mathbf{Y}_k, \mathbf{Z}_k) P_\ell^{\mu+1}-1}{\mu+1}\right] \cdot \exp\left[\varepsilon_{\ell ki}\right], \quad (29)$$

where

$$B_{ki}(\mathbf{Y}_k, \mathbf{Z}_k) = \exp\left[\bar{\alpha} + \xi z_i + \beta \mathcal{E}[B_y(\mathbf{Y}_k)] + \gamma \mathcal{E}[B_z(\mathbf{Z}_k)] + \nu_k + \eta_i\right], \quad (30)$$

$\delta > 0, \mu < 0, \mathbf{Y}_k$ and $\mathbf{Z}_k$ denote the vectors of individual i's neighbors' demand and of their demographic characteristics in cluster k, $\mathcal{E}[B_y(\mathbf{Z}_k)]$ and $\mathcal{E}[B_z(\mathbf{Z}_k)]$ denote scalar functions of neighbors' demand and of characteristics, respectively, with expectations being taken conditional on k, and preference parameters $\zeta_i, \bar{\alpha}, \xi, \beta, \gamma$ are unrestricted. The demand for housing services follows from (29) by Roy's identity. After taking logs we write and have for $y = \ell n Y$:

$$y_{\ell ki} = \mu P_\ell + \delta I_i + \bar{\alpha} + \xi z_i + \beta \mathcal{E}[B_y(\mathbf{Y}_k)] + \gamma \mathcal{E}[B_z(\mathbf{Z}_k)] + \nu_k + \eta_i.$$

I note that the slopes of the "indirect indifference curves" in (g_ℓ, P_ℓ) space (assuming that g_ℓ is a scalar):

$$-\frac{\frac{\partial V_{\ell ki}}{\partial_{g\ell}}}{\frac{\partial V_{\ell ki}}{\partial P_\ell}} = \frac{\zeta_i}{B_{ki} P_\ell^\mu},$$

are positive and increasing in price (given $\mu < 0$) and in the parameter ζ_i, that evaluates the tract-specific attributes $\mathbf{g}_\ell$. Therefore, other things being equal, tracts with better amenities are more attractive.[12] The term ν_k on the right hand side of (30) denotes an idiosyncratic characteristic of neighborhood cluster k, a random variable that is assumed to be independent and identically distributed across neighborhood clusters within each census tract. It is assumed to be unobservable to households when tract and cluster choices are made, but its value is revealed once households have chosen a particular cluster k. It is thus common among all households that reside in the same cluster. The term η_i is a random household taste parameter, that is observable by individual i but unobservable by the analyst; it is assumed to be independently and identically distributed over all individuals. The model yields, see below, that random variables (ν_k, η_i), which are unobservable by the analyst, make up the error component of the housing demand equation, with ν_k being a cluster-specific random effect in housing demand, and η_i an i.i.d. stochastic shock.

Note that the tract-specific term $\zeta_i \mathbf{g}_\ell$ in (28) can be specified to include tract characteristics interacted with individual characteristics. In this fashion, households with children may assign a different weight on school quality than a household with no

[12] This model is not designed to sustain equilibrium community formation, unlike Epple and Sieg (1999) and Epple, Romer and Sieg (2001). Therefore, it is not a drawback that it does not satisfy the single-crossing property with respect to income. This could, of course, be accommodated, but would make the error structure less transparent.

children, or households might value the presence of neighbors of the same ethnic background or race differently that those from other ethnic backgrounds or races.

Housing price P_ℓ, an index according to (27) that is homogeneous of degree one with respect to its components, is constructed as follows. The neighborhood amenities component $P_{\ell,nei}$, which *does vary* over tracts ℓ within each metropolitan area, is standardized relative to an overall mean for the entire economy. The price per unit of housing services component P_{stru}, that is associated with a dwelling's structure and does *not vary* across tracts within a metropolitan area (which can be justified by the assumption that the market for housing construction materials is competitive at the level of the metropolitan area), is standardized relative to an MSA-specific intercept. The price index P_ℓ expresses a key characteristic of housing markets, in that both components of the good "housing services" are bundled together.

Consistently with the price index, we define housing consumption to be the continuous flow of services that comes from the dwelling structure and neighborhood amenities. Empirically, dwelling structure is measured in terms of characteristics such as a dwelling's age, its number of bedrooms and baths, the availability of a garage and various structural quality features. Neighborhood amenities are proxied by the socioeconomic characteristics in the tract where a dwelling unit lies. Housing expenditure reflects both components of housing demand relative to nonhousing consumption.

The demand for housing structure, $y_{stru,\ell,k,i}$, conditional on neighborhood choice, follows from the conditional utility function $V_{\ell ki}$ using Roy's identity[13] with respect to price $P_{\ell,stru}$. After taking logarithms, this yields:

$$\begin{aligned} y_{stru,\ell ki} = \alpha + \vartheta p_{\ell,nei} + [\mu(1-\vartheta) - \vartheta] p_{stru} + \delta \ell n I_i + \xi z_i + \beta \mathcal{E}[B_y(\mathbf{y}_{stru,k})] \\ + \gamma \mathcal{E}[B_z(\mathbf{z}_k)] + \nu_k + \eta_i, \end{aligned} \tag{31}$$

with $\alpha \equiv \bar{\alpha} + \ell n(1-\vartheta)$, a parameter, lower case p's indicating the natural logarithm of the respective price variable, e.g., $p_{\ell,nei} = \ell n P_{\ell,nei}, p_{\ell,stru} = \ell n P_{\ell,stru}$.

Again, invoking the terminology of Manski (1993), we refer to the term B_y $(\mathbf{y}_k)$ on the right hand side of equation (31) as an *endogenous social effect:* a person's behavior depends on the actual *behavior* of her neighbors. I refer to the term B_z $(\mathbf{z}_k)$ as a *contextual effect*, a social effect which reflects taste over the characteristics of one's neighbors such as their race, ethnicity, and income. The unobserved stochastic components on the right-hand-side of equation (31) may reflect a conditional version of what Manski calls a *correlated effect:* similar individuals are likely to make similar choices of dwelling units and neighborhoods and therefore have unobserved characteristics in common.

[13] We are aware of an inconsistency here. This step requires that Roy's identity be taken with respect to a suitably restricted indirect utility function, where the appropriate neighborhood-specific components of the consumption bundle are held fixed, once neighborhood is chosen. However, this particular approach is not structural, in the econometric sense, and designed to highlight the complexities of housing as a joint discrete and continuous decision.

The logic of the endogenous social effect (a "keeping up with the Joneses" effect here) suggests an equation like (31) for each of the members of neighborhood cluster k. Thus, solving them simultaneously allows the analyst to obtain an instrument for $B_\gamma\ \mathbf{Y}_k$, the endogenous social effect on the right-hand-side of equation (31). This has implications for identification, to which we return in detail in Section 3.2.

3.1.2 Neighborhood choice

We assume that households limit their search to the metropolitan area in which they are observed to live. The probability that household i choose tract ℓ_i, from among tracts $\ell = 1, \ldots, \mathcal{L}$, and neighborhood cluster k_i, from among clusters $k = 1, \ldots, K_\ell$, is given by the probability that the (logarithm of) actual utility from this choice exceeds the utilities from all other choices:

$$Prob_{\ell_i k_i i} = Prob\{\ell n\omega_{\ell_i k_i i} - \ell n\omega_{\ell k i} + (\zeta_i g_{\ell_i} - \zeta_i g_\ell) \geq -(\varepsilon_{\ell_i k_i i} - \varepsilon_{\ell k i}); \forall(\ell, k) \neq (\ell_i, k_i)\}. \tag{32}$$

This can be computed, once the stochastic structure in equation (29) has been specified. It follows from equation (32) that when comparing utility between any two tracts the term $\frac{I_i^{1-\delta}-1}{1-\delta}$ cancels out. However, income and other individual characteristics are still present through the specification of $\zeta_i g_\ell$, the household-specific terms interacted with tract characteristics and the terms $\exp\left[-\frac{B_{ki}P_\ell^{\mu+1}-1}{\mu+1}\right]$, introduced in equation (29) above. In view of the definition of B_{ki} in (30) above, when comparing utilities across tracts, households are assumed to take expectations with respect to ν_k, which is assumed to be independent of other variables and unobservable at that point in the choice process. Therefore, the choice probabilities (32) are expressed as the probabilities of the events:

$$\left\{\zeta_i \mathbf{g}_{\ell_i} - \zeta_i \mathbf{g}_\ell - \frac{P_{\ell_i,stru}^{(1-\vartheta)(\mu+1)}}{\mu+1}\left(\widetilde{B}_{k_i i} P_{\ell_i,nei}^{\vartheta(\mu+1)} - \widetilde{B}_{ki} P_{\ell,nei}^{\vartheta(\mu+1)}\right) e^{\nu_k+\eta_i} \geq -(\varepsilon_{\ell_i k_i i} - \varepsilon_{\ell k i}); \right.$$
$$\left. \ell \neq \ell_i, k \neq k_i)\right\}, \tag{33}$$

where $\widetilde{B}_{ki} = \exp\left[\bar{\alpha} + \xi z_i + \beta\mathcal{E}[B_\gamma(\mathbf{Y}_k)] + \gamma\mathcal{E}[B_z(z_k)]\right]$.

Condition (33) has the intuitively appealing implication that the larger the value of the unobserved taste parameter η_i (the i.i.d. shock in the demand equation (31)), that is, the larger the dwelling a household wants given all observables, the smaller the neighborhood price it wishes to pay. Therefore, variation of price across tracts, as expressed by component P_ℓ in the composite price index (27), is a key element of the interaction between the discrete choice of neighborhood and the continuous choice of housing structure.[14]

Under the assumption that the $\varepsilon_{\ell ki}$'s in equations (32) and (33) are independently and identically extreme-value distributed across all census tracts in an MSA and in all neighborhood clusters within them, the choice probabilities are given by the multinomial logit model (MNL):

$$\mathrm{Prob}_{\ell_i k_i i} = \frac{\omega_{\ell_i k_i i} e^{\zeta_i g_{\ell_i}}}{\sum_{\ell=1}^{\mathcal{L}} \sum_{k=1}^{K_\ell} \omega_{\ell k i} e^{\zeta_i g_\ell}}. \tag{34}$$

There is a well-known drawback to applying the MNL model, in that it stretches the plausibility of the stochastic structure, as the $\varepsilon_{\ell ki}$'s are unlikely to be independent across alternative residential choices. In particular, evaluating alternative clusters within the same census tract will involve common tract-level unobservables that will cause the error terms to be correlated. One can fully account for this possibility via a nested logit model or more general models.[15]

In view of lack of information about which neighborhoods (either at the cluster- or tract-level) households considered before they choose to locate where they are observed, Ioannides and Zabel assume individuals choose over Census tracts. They utilize a suggestion of McFadden (1978)[16] that the discrete choice model may be estimated by generating a random sample of alternatives from the full choice set, which may be unobserved.[17]

For each dwelling unit observed in the cluster subsample of the public NAHS data Ioannides and Zabel identify, by relying on *confidential* U.S. Census data, the Census tract in which it lies and choose randomly ten other tracts from among the universe of tracts in the respective metropolitan area. They approximate the expressions in the choice probabilities above by using as regressors tract-level characteristics, on their own and also interacted with individual characteristics. They also include individual

[14] This interaction between η_h and tract specific variables introduces heteroscedasticity in the errors of the discrete choice problem.

[15] The expected value of the maximum utility associated with the neighborhood choice problem (32), when the $\varepsilon_{\ell ki}$'s are assumed to obey a Generalized Extreme Value (GEV) distribution [McFadden (1978)], may be written as: $\mathcal{E}\left\{\max_{k\in K_\ell, \ell\in\mathcal{L}} \Omega_{\ell ki}\right\} = \sum_{\ell=1}^{\mathcal{L}}\sum_{k=1}^{K_\ell} \Omega_{Lki} \cdot \mathrm{Prob}_{\ell_i k_i i} = \ell n G\left[(\Omega_{11i})^{-1}, \ldots, (\Omega_{\ell ki})^{-1}, \ldots, (\Omega_{\mathcal{L}K_{\mathcal{L}}i})^{-1}\right]$, (35) where the function $G[\cdot]$ satisfies McFadden's conditions for a generating function for a GEV distribution $F(\varepsilon) = e^{-G}[e^{-\varepsilon_1}, \ldots, e^{-\varepsilon_N}]$. See Ioannides (2010), Chapter 3, Appendix B, for more details.

[16] See Blackley and Ondrich [3] and Quigley [25] for two exceptions, and Fox (2007) for a modernization of McFadden's procedure. Bierlaire *et al.* (2003) shows that the consistency of estimation when using a subset of the opportunity set extends to all random utility models where the errors obey a GEV distribution.

[17] This estimation is consistent, provided that: one, independence from irrelevant alternatives holds; and two, if an alternative is included in the assigned set, then it has the logical possibility of being an observed choice from that set. The first condition is ensured by their use of the MNL; the second is satisfied because random selection satisfies the "uniform conditioning property" of McFadden, *op. cit.*, 88–89. See also Fox (2007).

variables interacted with statistics of the joint distributions of tract-level variables to proxy for the inclusive value, an auxiliary function that is included in the second stage of the nested logit model and which captures the heterogeneity of clusters within the tract. See Ioannides and Zabel, *op. cit.*, Appendix A. Their approach falls short of a full structural estimation of the discrete choice model, but does allow the data to determine which particular statistics, say from among different statistics interacted with individual characteristics, best explain the observed choices.

3.2 Housing demand with neighborhood effects

The conditional demand for housing structure equation by individual i in metropolitan area m, census track ℓ, cluster k, is:

$$y_{stru,m\ell ki} = \alpha + \vartheta p_{m,\ell,nei} + \vartheta' p_{m,stru} + \delta \ell n I_i + \beta \mathcal{E}\left[y_{stru,k(i)}\right] + \gamma \mathcal{E}\left[z_{k(i)}\right] + \nu_k + E[\eta_i|\ell = \ell_i] + \psi_i, \tag{36}$$

where $\vartheta' \equiv \mu(1 - \vartheta) - \vartheta$, a parameter, and $k(i)$ denotes the neighborhood cluster of individual i. Note that the endogenous and contextual effects have been specified as the means of the neighbors' housing demand and of a vector of neighbor characteristics; that is $B_y(\mathbf{Y}_k) = \mathcal{E}\left[y_{k(i)}\right]$, and $B_z(\mathbf{z}_k) = \mathcal{E}\left[z_{k(i)}\right]$, respectively.

The conditional mean correction in equation (36), $E[\eta_i|\ell = \ell_i]$, accounts for the fact that the error term on the right-hand-side of the demand equation (31) is likely to be correlated with the other regressors in the model. This correction term can be estimated using the results from the neighborhood choice equation (34) in the standard fashion for sample selection bias.[18] While inclusion of the components of the price index $(p_{m,stru}, p_{m,\ell,nei})$, readily follows from the model, allowing unconstrained estimation of their coefficients may determine whether neighborhood demand is a substitute or a complement to housing structure demand.

The mean of the neighbors' housing demand is correlated with the error term in equation (36) since it includes the unobserved cluster effect, ν_k. In order to identify the model, given the reduced form, one needs an instrument, a variable whose neighborhood average is not included in the causal model (e.g., is *not* a contextual effect). As shown by Brock and Durlauf (2001), the selection correction terms are valid instruments. In particular, the neighborhood averages of these terms emerge naturally, from solving endogenously for the expectations of neighbors' demands, as valid instruments.

Identification is an issue even in the absence of social interactions. In order to identify the housing demand model that is conditional on residential choice, the selection

[18] The specific form of the sample selection bias correction terms are computed using the results in Dubin and McFadden (1984). The correction requires eleven terms, one for each of the eleven census tracts in the neighborhood choice model. See Ioannides and Zabel (2008), Appendix B, for details.

correction terms, (making up $E[\eta_i|\ell = \ell_i]$), must not be collinear with the other regressors in equation (36). One way of achieving identification is by ensuring that one or more variables in the neighborhood choice model be excluded from the housing structure demand equation. Given that we are modeling housing structure demand, there are variables that affect neighborhood choice and not structure demand and thus qualify. Thus, the housing structure demand equation is identified via these exclusion restrictions.

The identification of the endogenous neighborhood effect, the mean of neighbors' housing demand, $\mathcal{E}[y_{k(i)}]$, rests on solving the housing structure demand equations (in expectation) for all members of a cluster as a simultaneous system for $\mathcal{E}\left[y_{stru,k(i)}\right]$. Thus, the cluster means of the sample selection terms, $\mathcal{E}[\eta_i|\ell = \ell_i]$, arise naturally as identifying instruments:

$$\mathcal{E}[y_{stru,k(i)}] = \pi_0 + \pi_1 p_{mk_i,nei} + \pi_2 p_{m,stru} + \pi_2 \mathcal{E}[z_{k(i)}] + \pi_3 \mathcal{E}[\eta_h|s = s_h] + \eta_{k(i)}, \qquad (37)$$

where $\eta_{k(i)}$ is the unobserved error term and the neighbors' mean income $\overline{I_{k(i)}}$ is included in $\mathcal{E}[z_{k(i)}]$ for brevity.

The eleven sample selection bias correction terms *vary within* the cluster, as they depend on individual variables via the interaction terms in the neighborhood choice model.[19] They are, therefore, key to the identification of the endogenous neighborhood effect. Intuitively, one's neighbors' selection bias correction terms are excluded from one's own demand for housing structure equation because one's neighbors' tastes for housing (in contrast to their observed characteristics) do not directly affect one's own demand for housing. These preferences do have an indirect effect, though, through the endogenous neighborhood effect.

3.3 Estimation of neighborhood choice and housing structure demand with neighborhood effects

Ioannides and Zabel (2008) construct measures for housing price and quantity, both quantities that are not are directly observable. They decompose continuous housing demand into two components; structure demand and neighborhood demand. They model the former component in terms of a continuous scalar quantity that represents the flow of housing services. The price of housing is thus the price for a unit of services from housing structure. Neighborhood demand may be either a substitute or a complement to housing structure, and including the neighborhood price in the demand for structure equation allows for the data to resolve the issue.

[19] This feature addresses the critique of Angrist and Pischke (2009) of specification of peer effects, Section 4.6.2, p. 192–197, in that the equation accounts for variation in ex ante peer characteristics that predate the outcome variable, as in the determinants of neighborhood choice.

Ioannides and Zabel estimate an ad hoc hedonic house price function, $P_{m\ell ht}$, for a dwelling unit h as a function of its structural characteristics $\mathbf{x}_{m\ell ht}$, location in MSA m, and in census track ℓ with characteristics $\mathbf{g}_{m\ell t}$, using the noncluster NAHS data at time t. The use of these data has two advantages. One is that they make up approximately ninety per cent of the NAHS data thus affording them a much larger data set than the neighborhood clusters subsample of the NAHS. Second, the prices thus obtained come from a different data set than the one used to estimate the housing structure demand equation. Hence, they use the tract characteristics to proxy for all levels of neighborhood quality and, in particular, of the consumption of housing structure by a unit's neighbors and their socioeconomic characteristics, $(\mathbf{Y}_{k(i)}, \mathbf{z}_{k(i)})$.

Their hedonic equation is specified as follows:

$$\ell n P_{m,\ell,h,t} = \sum_{m=1}^{M} a_{0mt} MSA_{m,h,t} + a_1 \mathbf{x}_{m,\ell,h,t} + a_2 \mathbf{g}_{m,\ell,t} + u_{m,\ell,h,t}, \tag{38}$$

where h indexes dwelling units, $i = 1, \ldots, N_m, \ell = 1, \ldots, \mathcal{L}_m$, census tracks in MSA m, and $m = 1, \ldots, M$ MSAs for each of the same three waves of the NAHS data for which data on clusters are also available, $t = 1985, 1989, 1993$; MSA_{mht} is dummy variable equal to 1, if unit h is in metro area m in period t, and equal to 0, otherwise. Based on (38), Zabel (2004) defines a price index for the average (structure) quality dwelling unit in tract ℓ in MSA m and wave t as:

$$P_{m,\ell,t} = \frac{\exp\left(\hat{a}_{0mt} + \hat{a}_1 \bar{\mathbf{x}} + \hat{a}_2 \mathbf{g}_{m\ell t}\right)}{\exp\left(\hat{a}_{011} + \hat{a}_1 \bar{\mathbf{x}} + \hat{a}_2 \bar{\mathbf{g}}\right)}, \tag{39}$$

where the index is relative to MSA code = 1 in time period 1 (that is arbitrarily standardized relative to Denver in 1985), and q and g are evaluated at fixed-mean values, $\bar{x}$ and $\bar{g}$, respectively. Note that $p_{111} = 1$. This price index is decomposed into:

$$P_{m,\ell,t} = \frac{\exp\left(\hat{a}_{0mt} + \hat{a}_1 \bar{\mathbf{x}}\right)}{\exp\left(\hat{a}_{011} + \hat{a}_1 \bar{\mathbf{x}}\right)} \cdot \frac{\exp\left(\hat{a}_2 \mathbf{x}_{m\ell t}\right)}{\exp\left(\hat{a}_2 \bar{\mathbf{g}}\right)} = P_{m,t,stru} \cdot P_{m,\ell,t,nei}, \tag{40}$$

where $P_{m,t,stru}$, the component of price that corresponds to structure and is invariant within MSA m, and $P_{m,\ell,t,nei}$ that corresponds to neighborhood (tract) ℓ. The neighborhood price varies across census tracts. The prices will be referred to in logs, $p = \ell n P$.

Housing structure services for an individual i living in unit $h(i)$ are defined, in logs, as

$$y_{m\ell ki,stru} = \ell n r + \hat{a}_{0mt} + \hat{a}_1 \mathbf{x}_{m\ell ht} - P_{m,t,stru}. \tag{41}$$

Once the hedonic equation (38) is estimated, the demand for structure, according to equation (41), and the structure and neighborhood prices $p_{m,t,stru}$, $p_{m,\ell,t,nei}$ according to equation (40), can be computed.

3.3.1 Estimation results of neighborhood choice

Ioannides and Zabel estimate the model of neighborhood choice according to Equation (34). The structural characteristics of dwelling unit m, ℓ, h, t, $\mathbf{x}_{m\ell ht}$, include the age of the unit and its square, the number of full baths, of bedrooms, and of total rooms, whether or not there is a garage and a number of additional structural quality variables (such as whether the enumerator saw cracks on walls and ceilings, broken pipes, etc.). The neighborhood characteristics of tract m, ℓ, h, t, $\mathbf{g}_{m\ell t}$, include a dummy variable that indicates whether or not the unit lies in the central city of the MSA, the property tax rate, and tract-level variables that include median household income, the percent over 25 years of age who graduated from high school, and the percent of the tract population that is nonwhite.

A benchmark model that contains only tract-specific characteristics of individuals and dwellings in the tract of the current residence shows that higher price, median age of dwellings, median number of bedrooms, the fractions of owners and nonwhites in the tract and the fractions of residents in the tract with a high school degree and commuting less than twenty minutes all increase the likelihood of choosing a tract. On the other hand, a higher median income, median rent, median age of tract residents, vacancy rate, poverty rate, and unemployment rate decrease the likelihood of choosing a tract. The fraction moved in within the last 5 years is not significant. This regression has a pseudo-R^2 of 0.0736. The results show that the valuation of median tract income is increasing in individual income and that the valuation of vacancy rates declines with income though it increases in an absolute sense. On the other hand, while individuals positively value homeownership rates, this valuation declines with income.

Interacting the tract-level race variables with the individual race variables shows that an increase in the percent nonwhite will decrease the likelihood of tract choice but there is no additional effect if there is at least 50% nonwhites in the tract (this is measured through the variable dominant race). For nonwhites, an increase in the percent nonwhite will increase the likelihood of tract choice. There is an additional positive effect if there is at least 50% nonwhites in the tract.

Interacting discrete indicators for individual education, that is, for those without a high school degree, with at most a high school degree, and with a college degree, with the fraction of individuals over 25 years old in the tract who have a high school degree shows a strong positive relationship between individual and tract education; an increase in the fraction in the tract with a high school degree will decrease the probability of residing in the tract for individuals with no high school degree and will increase this probability for those with a college degree. Similarly, results with the median age shows that increasing the median age makes those in the first quartile of the age distribution less likely to choose the tract compared to older individuals. It also makes married household heads less attracted by the tract. When a full complement of 74 explanatory variables are included, the pseudo-R^2 rises to 0.1628, and the additional variables included compared to those in the model reported in column 3 of Table 4 are jointly statistically significant.

3.3.2 Estimation results for housing structure demand with neighborhood effects

Table 5, *ibid.*, reports the estimation results for the housing structure demand equation (36), measured in logs. The regressors include the logs of the structure and neighborhood prices, and in addition the log of income, the number of persons in the household, and dummy variables that indicate if the owner has graduated from high school, is married, is white, and moved in the last five years as individual variables and their cluster averages as contextual effects, and the endogenous social effect.

When the eleven sample selection bias correction terms are included, they are statistically significant as a group at the 1% level. The estimated price elasticities for structure and neighborhood, respectively, are −0.1784 and 0.2086. The signs suggest that structure and neighborhood quality are substitutes. The permanent income elasticity is positive and significant; 0.2106. Household size has a positive and significant effect on the demand for housing structure. The instrumental variable estimates of demand for structure show that the structure price elasticity is now much smaller than when the neighborhood effects were not included (and not significant). The neighborhood price elasticity is now negative, small in magnitude, but only marginally significant. Clearly, the neighborhood price is positively correlated with the neighborhood effects and hence there is a positive bias when the latter are excluded from the demand equation. The negative coefficient for the neighborhood price indicates that structure and neighborhood are complements. The income elasticity is also much smaller in magnitude, though it is still positive and significant. The coefficient estimate for the mean of the neighbors' structure demand, the endogenous social effect, is 0.8504. The contextual effects are not jointly significant and only one variable is individually marginally significant; the mean of neighbors' household size (p–value = 0.048). The sample selection bias correction terms are marginally significant as a group (p–value = 0.041). These terms are more significant in the regressions without neighborhood effects. This is also true for the price and income elasticities. Clearly, these terms are picking up some of the omitted neighborhood effects.

Overall, the estimates of the housing structure demand therefore do confirm that endogenous neighborhood effects are important and are in fact *strengthened* when neighborhood choice is accounted for. Further, the own-price elasticity nearly doubles in magnitude (−0.1319 versus −0.0772) but remains insignificant. It is also true, however, that the unobservable effect for individuals in the same neighborhood is an important part of the story, even after neighborhood choice has been accounted for.

3.4 Neighborhood information in hedonic regressions

In view of these empirical results, we can reflect on the results of Kiel and Zabel (2009), who use the same data as Ioannides and Zabel (2008), and discuss empirical hedonic results with social effects. These authors estimate housing hedonics with neighborhood information that correspond neatly to the Ioannides and Zabel model and empirical approach. They estimate a hedonic function $\Pi(\mathbf{x}_h, \mathbf{x}_k, \mathbf{g}_\ell; \mathbf{z}_k; \mathbf{z}_{\ell(k)})$, whose arguments are attributes of the dwelling, cluster, tract, $\mathbf{x}_h, \mathbf{x}_k, \mathbf{g}_\ell$, and contextual effects associated with the occupants of the dwelling units in the cluster and tract of unit $h, (\mathbf{z}_k, \mathbf{z}_{\ell(k)})$, while accounting for a base MSA-specific price. Their results are obtained with cluster random effects and robust standard errors and generally confirm the notion that cluster, tract and MSA variables ("location, location, location," 3L's in their words) are all highly significant in the house price hedonic. When the attributes of these different aspects of location are alternatively excluded from the regression, the percent increase in the standard error are quite similar: 2.2%, 2.3%, and 2.7%, respectively. This indicates that each of the 3Ls have a similar importance in determining house prices. Kiel and Zabel report on results when data on clusters are excluded, which are after all a very special feature of the NAHS. Doing so does not affect much the estimates for the coefficients of dwelling attributes, nor for those for the census tract attributes. Yet, it is particularly noteworthy that the coefficient of the MSA-specific price increases from 0.677 to 1.022. This indicates that the cluster variables are important in housing values and relevant for the construction of house indices.

These results also suggest that the concept of neighborhood is multifaceted. Individuals indeed care about the quality of neighborhoods at several levels (scales). The information at the levels of cluster, tract and MSA can be highly correlated, but Kiel and Zabel suggest that there is also independent information at those different levels that appears to have significant impact on the willingness to pay for a house in a given location. This accords with the notion that different small neighborhoods have different character, and their proximity to one another confer character to higher-level neighborhoods that contain them.

Finally, while the use of the hedonic price function here is empirically well grounded, recalling the generalization of the Nesheim model in Section 2.6 suggests that the hedonic estimation itself also contributes to estimation of structural parameters. This is the import of the Ekeland, Heckman and Nesheim (2004) critique of standard hedonic estimations. Ioannides (2010), Chapter 3, Appendix C, proposes a conceptual extension of the modern hedonic approach of Ekeland *et al.* to housing in the presence of contextual and endogenous neighborhood effects and briefly discusses the pitfalls of misspecification. The multitude of demands imposed on the existence of hedonic price functions makes it difficult to obtain in closed form hedonic price functions for arbitrary preferences. Nonetheless, as we discussed in Section 2.8, Bayer and Ross, *op. cit.*, demonstrate a notable use of the monotonic but highly nonlinear relationship between neighborhood housing price and neighborhood quality as an empirical control function for neighborhood unobservables.

4. NEIGHBORHOOD EFFECTS AND THE GEOMETRY OF THE CANONICAL URBAN MODEL

Different communities differ in terms of density of settlement, quality of amenities and of public services, distance to employment centers and to transportation systems. Variations also exist within different parts of larger communities. Such differences are important in determining the fabric of large cities and metropolitan areas more generally. When individuals choose where to live, they take into consideration many such attributes of neighborhoods in addition to specific features of the dwelling units that they choose to live in. Unlike the models developed above, here we impose a specific geometry. Individuals commute to a central business district (CBD) in order to work and/or socialize.

We assume that the typical individual derives utility from consuming housing in quantity $h(\ell)$ and a composite nonhousing good, $c(\ell)$, when located in location ℓ, indicated by the distance from a city's *predetermined* central business district (CBD) along a line.[20] The canonical urban model implies stark income or preference segregation. This property of the model provides a natural benchmark against which to gauge the impact of neighborhood effects. We consider neighborhood effects indirectly through the impact on location decisions of individuals of exogenous and endogenous amenities. In particular, exogenous amenities valued by individuals may cause high-income individuals to live closer to, or further away from, the CBD, depending upon the strength of these preferences. Such patterns must not be necessarily attributed to high-income individuals' wanting to be near other high-income individuals. Endogenous amenities valued by individuals, in the form of individuals' valuing being near other individuals with particular characteristics, generate a richer set of possibilities, including multiple equilibria. One could speculate, in view of the results of Brock and Durlauf (2007) discussed in Section 2.4, that pattern reversals that are due to multiple equilibria, may be used to identify presence of endogenous social effects in location patterns at the city level. Brock and Durlauf demonstrate that under various shape restrictions on the probability density of the unobservables, pattern reversals can only occur because of multiple equilibria and hence endogenous social interactions. This idea has not yet been utilized in urban research.

Below we start with a presentation of the canonical urban model, also known as Alonso–Mills–Muth model. Next we use it to study the impact of amenities, first when their distribution over space is given exogenously and then when they are created by

[20] While the notion of a circular city is arguably more realistically appealing, a number of results have underscored that analytically there is no advantage to the circular configuration. In particular, as Ogawa and Fujita (1980; 1989) show, nonmonocentric urban configurations in two-dimensional space (with circular symmetry imposed) are qualitatively essentially the same as those in the one-dimensional space.

individuals' location decisions. We then turn to examination of the impact of income and preference heterogeneity on location decisions. We show that across locationally defined (that is, in terms of distance from the CBD) neighborhoods, income and housing preferences are positively correlated, but within neighborhoods, they are negatively correlated.

4.1 The Alonso – Mills – Muth canonical urban model

Utility is denoted by $\Omega(h(\ell), c(\ell))$. Let $R(\ell)$ denote the rental rate of land. In the simplest possible case, we assume that housing consumption is simply consumption of services from land. The respective budget constraint is:

$$c(\ell) + R(\ell)h(\ell) = I - k(\ell), \tag{42}$$

where I denotes income, $k(\ell)$ transportation costs as a function of distance from the CBD, and the price of the composite good is assumed to be invariant to distance from the CBD and thus set equal to 1.

It is standard to start by considering the choice of (h, c), conditional on ℓ, so as to maximize utility $\Omega(h, c)$, subject (42), for which the first order condition is:

$$\frac{\Omega_h(h, c)}{\Omega_c(h, c)} = R(\ell). \tag{43}$$

Demand for housing follows from the solution to this problem.

It is convenient in developing the locational equilibrium model to work also with the dual representation of preferences. By solving for the optimal bundle $(h^*(\ell), c^*(\ell))$ and by substituting back into the utility function we obtain the indirect utility function,

$$\mathcal{O}(R(\ell); I - k(\ell)) \equiv \Omega(h^*(\ell), c^*(\ell)).$$

By requiring invariance with respect to distance from the CBD, we obtain a condition that the land rental rate must satisfy so as the typical individual be indifferent across all locations. That is, moving away from the CBD increases transportation costs and therefore land rent must decline to make it attractive to be there:

$$R'(\ell) = k'(\ell)\frac{\mathcal{O}_2}{\mathcal{O}_1}. \tag{44}$$

By integrating (44) with respect to ℓ, we get an expression for the equilibrium land rental rate, which of course, contains a constant of integration. Equilibrium conditions, which we take up in the next section, determine the model fully. Condition (44) is known as the Mills-Muth condition. The basic framework outlined so far combines features of Alonso (1964), Mills (1967) and Muth (1969).

An alternative approach is to work with the *bid rent function* [Fujita (1989)], defined as the maximum land rental an individual is willing to pay, conditional on attaining a given level of utility, say ω. That is,

$$\Psi(\ell, \omega = \max_{c,h} \left\{ \frac{I - k(\ell) - c}{h} | \Omega(h, c) = \omega \right\}. \tag{45}$$

It follows from this definition of the expenditure function that its value when the land rental rate is set equal to the bid rent function is equal to income net of transport costs. The bid rent function is decreasing in distance from the CBD and in utility attained. That is, a household can attain a higher utility level given income net of commuting costs only if land rent is reduced. Fujita (1989), p. 22, proves that if the transport cost function is linear or concave, the bid rent curves are strictly convex.

4.2 Locational equilibrium for individuals

Closing the model requires that we characterize equilibrium in the land market. Let land extend along a line away from the CBD. Expressing land supply equal land demand closes the model. Alternatively, we may assume that equilibrium utility is given from other opportunities in the economy. Setting equilibrium utility in the city be equal to what is exogenously given determines the constant of integration.

An example clarifies things. Let $\Omega(h, c) \equiv (1-\beta)\ell nc + \beta \ell nh$, for which $\mathcal{O} = \beta \ell n\beta + (1-\beta)\ell n(1-\beta) + \ell n(I - k\ell) - \beta \ell nR$. This yields $R(\ell) = R(0)(I - k(\ell))^{\frac{1}{\beta}}$. If the supply of land is given, then equating it with demand for land determines $R(0)$. Alternatively, if the opportunity cost of land is given, then by equating the land rental at the edge of the city to it yields a condition that $R(0)$ and the land size of the city, $\bar{\ell}$ must satisfy. So, by equating demand and supply of land all unknowns are determined. Finally, equating $\mathcal{O}$ from above to exogenously given utility determines directly the land rental rate and therefore from the condition for land equilibrium, the land size of the city. Equivalently, we may work with city size in terms of population.

As another example, consider that housing (that is, land) consumption is fixed at the unit level, so that for locational equilibrium, direct utility of consumption, $\Omega(1, I - k(\ell) - R(\ell))$, be constant for all ℓ. This yields a land rental rate that varies with $-k(\ell)$, that is it declines as distance from the CBD. We conclude that when distance from a given center is the sole index of locational choice and determinant of the transportation cost, then the equilibrium land rental fully reflects all relevant parameters.

This approach allows us to characterize locational choice of heterogeneous households. Given a land rental function, an individual locates so as to maximize utility, which occurs where the bid rent function is tangent to the land rental function. Provided that preferences are well behaved, transportation costs are increasing in the distance from the CBD and the Marshallian demand for land is increasing in income, one can prove that given two households that are otherwise identical, face the same transport cost function and differ in terms of incomes, then households with higher incomes locate further away from the CBD than households with lower incomes [Fujita (1989), 105–111]. With many household types, $j = 1, \ldots, J$, equilibrium may be characterized easily provided that bid rent functions corresponding to different types may be ranked in decreasing order of steepness.

The group with the overall steepest bid rent curve locates nearest the center, and then one with the next steepest follows. These results with the equilibrium land rental function being the upper envelope of the equilibrium locations of different groups are clarified in terms of the *boundary rent curve*, $\hat{R}_j(\ell)$. For every household type, this gives the rent at the edge of a zone occupied by the particular group so that all individuals of that type have been accommodated, given a particular utility level [*ibid*., p. 55]. Then, at equilibrium, the utility level the group may enjoy is determined by the value of the boundary rent curve associated with the next group. Locational equilibrium is determined recursively from the urban fringe. Given the value of land in agriculture, R_A, the boundary rent curve for the group occupying the outermost area yields the size of the urban area and the associated utility is given by the utility level that makes the bid rent curve at the respective location be equal to $R_A = \hat{R}_j(\ell_J)$, and $\Psi(\ell_J, u_J^*) = R_A$. Thereafter, $\Psi(R_j^*, u_j^*) = \hat{R}_j(\ell_j), j = 1, \ldots, J-1$. In the special case where households differ with respect to income only, and $I_1 < \ldots < I_J$, then the steepness of the bid rent curves varies inversely with income. As a result, concentric zones around the CBD will be occupied by groups in increasing order of income.

When households differ in terms of incomes, the system of land rentals that equilibrates the housing market reflects the entire distribution of incomes. Alternatively, given individuals who differ in terms of incomes, then the price system expresses how individuals in an urban economy self-organize vis a vis location. Land rentals thus reflect the socioeconomic characteristics of the entire population of an urban area.

4.3 Amenities

A central feature of the urban location model as presented so far is that the CBD is assumed to function as an amenity. The amenity role of the CBD provides an essential metric for the value of distance from it. Brueckner, Thisse and Zenou (1999) in an effort to analyze the role of amenities as distinct from travel to the CBD introduce an additional argument of the utility function, location-specific amenities denoted by, $a(\ell)$,

$$\Omega \equiv \Omega(h(\ell), c(\ell); a(\ell)).$$

Amenities may be exogenous or endogenous, that is, they may be exogenous attributes of locations, or endogenous being defined in terms of decisions that agents themselves make or of characteristics of the agents who choose to locate at different locations. The definition of the indirect utility function is adapted in the obvious way: $\mathcal{O}(R(\ell); I - k(\ell);\ a(\ell)) \equiv \Omega(h^*(\ell), c^*(\ell); a(\ell))$. With only a single type of households in the urban economy, locational equilibrium implies:

$$R'(\ell) = -\frac{k'(\ell)}{h(\ell)} + \frac{\mathcal{O}_3}{h(\ell)\mathcal{O}_2} a'(\ell). \tag{46}$$

Even if the amenity has a positive effect on utility, $\mathcal{O}_3 > 0$, the amenity must improve with distance from the CBD. Specifically, $a'(\ell)$ must be positive and sufficiently large for the land gradient to remain negative.

If two groups of individuals who differ only with respect to income, $I_1 < I_2$, compete for land, then in the absence of the amenity effect, each group will locate in the area of the city where it outbids the other group. Let $\hat{\ell}$ denote the threshold point, that is the location where the land bid rent curve for the two groups become equal: $R_1(\hat{\ell}) = R_2(\hat{\ell})$. At the threshold point, both group face the same price. So, the difference in housing consumption is due to the difference in income net of transportation costs. It is reasonable to assume that transportation costs are higher for the higher income group. So: $h_2(\hat{\ell}) > h_1(\hat{\ell})$. Under the assumption that the unit cost of transportation depend on the opportunity cost of time, then $k_1 < k_2$. The location of the two groups depends on the difference between the slopes of the bid-rent curves at $\hat{\ell}, R_2'(\hat{\ell}) - R_1'(\hat{\ell})$. That is, from (46) this difference is equal to:

$$-\left(\frac{k_2'\hat{\ell}}{h_2(\hat{\ell})} - \frac{k_1'\hat{\ell}}{h_1(\hat{\ell})}\right) + a'(\hat{\ell})\left(\frac{\mathcal{O}_3(R_2(\hat{\ell}); I_2 - k_2(\hat{\ell}); a(\hat{\ell}))}{\mathcal{O}_2 h_2(\hat{\ell})} - \frac{\mathcal{O}_3(R_1(\hat{\ell}); I_1 - k_1(\hat{\ell}); a(\hat{\ell}))}{\mathcal{O}_2 h_1(\hat{\ell})}\right). \tag{47}$$

Ignoring for a moment the dependence of indirect utility on amenities, whether or not the higher income group will occupy the land nearer the CBD depends on the different effect of income on the demand for housing: if housing demand rises less rapidly relative to unit transportation costs as income increases, then $\frac{k_2'(\hat{\ell})}{h_2(\hat{\ell})} > \frac{k_1'(\hat{\ell})}{h_1(\hat{\ell})}$, the rich live near the CBD and the poor in the suburbs; if, on the other hand, unit transportation costs rises less rapidly relative to housing demand as income increases, $\frac{k_2}{h_2(\hat{\ell})} < \frac{k_1}{h_1(\hat{\ell})}$, then the rich live in the suburbs and the poor near the CBD. To obtain a prediction for land use in the presence of the amenity, we need to assess how the term $\frac{\mathcal{O}_3(\ell)}{h(\ell)}$ varies with income. Brueckner *et al.* show that if the (constant) elasticity of substitution between nonhousing consumption, housing consumption and the services of the amenity is less than 1, then the above ratio increases with income. If the amenity declines with distance from the CBD but its marginal effect with distance is small, then the poor will live in the center and the rich in the suburbs. If, on the other hand, the marginal effect with distance is negative but large in absolutely value, then the amenity advantage of the center will draw the rich near the CBD and the poor to the suburbs. So, increasing the amenity attractiveness of the CBD may reverse the location pattern.

These results are critical for understanding the impact of income differences on land use in the presence of neighborhood effects as a "centrifugal force," that is the attraction of the CBD, even when it is endogenous. The amenity has so far been modeled as dispersed according to a exogenous pattern. Nonetheless, its impact on the geometry of land use is particularly useful in modeling the impact of correlated effects on urban land use at equilibrium. That is, we may express a number of assumptions about preferences for housing and the amenity, and their variation with distance from the CBD and explore their impact on equilibrium land use.

Another use of the above model is to explore the impact of endogenizing the amenity. We may assume that individuals differ and define the amenity so as to be a function of the characteristics of one's neighbors relative to one's one characteristic. Let income be such a characteristic. If individuals prefer to live near others with similar incomes, then that would be so wherever they might be. If a rich person were to move into a poor area, then her bid-rent curve would be depressed because she would have to live near others who are not like her. But, rich people moving together near the CBD may cause an increase of each other's bid-rent curves and thus force a reversal of the location pattern. However, the endogenous amenity effect must be sufficiently strong to accomplish this. We explore this in more detail next.

4.3.1 Endogenous amenity

As we saw, individuals' valuing just being near others defines a metric that implies in turn a density, or equivalently, a bid-rent curve. What if individuals differ with respect to income, an exogenous characteristic that is salient with respect to housing decisions? While income is exogenous, where individuals locate at equilibrium is endogenous and plays the role of an endogenous amenity of the type analyzed by Brueckner, *et al.*, *op. cit.*.

These authors simplify the problem by considering two types of dwellings only, $\bar{h}_1$ and $\bar{h}_2$, $\bar{h}_2 > \bar{h}_1$, with both being available in each location and with group 2, the rich, choosing the larger size. They show that for the pattern 'rich in center' and 'poor in suburbs' to be an equilibrium, the bid-rent by the poor, although steeper than those by the rich within the either zone, must be entirely dominated by that of the rich in center. This is compatible with a discontinuity at the boundary. Such discontinuity is of course an artifact of the endogenous amenity's being insensitive to how far one is from the boundary. Such a discontinuity gives rise to the possibility that a mixed city is an equilibrium outcome. That is, the bid-rent of the rich, when the endogenous amenity is the mean neighborhood income, would have to dominate that of the poor over the entire domain, so that a city may accommodate both groups throughout.

This condition is more stringent, in that the rich bid-rent would in this case be below that of the rich alone case. However, the outcome is disadvantageous for the poor, in that they would have to pay a higher rent than they would if they were to occupy the center on their own. Such a disadvantageous desegregation is a new possibility within the urban model and whose additional properties, including stability, need to be explored further. See Brueckner *et al.* (1999) for full details.

4.4 Households with heterogeneous preferences

Critical to understanding the impact of neighborhood effects is to be able to compare with equilibrium outcomes when individual differ but there are no neighborhood effects. Individuals may differ in terms of preferences or income, or preferences may covary with incomes. We examine in turn the impact of preference heterogeneity in the form of additive shocks to utility on location decisions of households and on equilibrium in the canonical model, in terms of income, and finally in terms of both preferences and income.

4.4.1 Preference heterogeneity

We consider next taste heterogeneity in the form of random shocks to utility. Following the approach of Anas (1990), let locations be defined in terms of a denumerable number of thin areas which are indexed by their distance from the CBD, $\ell = 0, \ell_1, \ldots, \ell_{N-1}$, with transport costs being negligible within each location and being equal to k, per unit distance, across locations. All individuals commute to the CBD. Let an individual i who considers locating in location ℓ derive utility

$$\widetilde{\Omega}_{i\ell} = \mathcal{O}(R(\ell); I - k\ell) + \varpi_{i\ell}, \ell = 0, 1, \ldots, N-1, \tag{48}$$

where the random variables $\varpi_{i\ell}$ are independently and identically distributed across all individuals and locations. Essentially, individuals are ex ante identical. However, individual i, $i \in \mathcal{I}$, in evaluating each site ℓ draws from a distribution, which is the same for all sites, a component that is unobservable to the analyst. Assuming that the distribution of the $\varpi_{i\ell}$'s is extreme value of type II, with mean zero, variance $\frac{\pi^2}{6\varrho^2}$ and mode $-\frac{EC}{\varrho}$, where EC $= 0.5772$, Euler's constant, the resulting choice probabilities are given by:

$$\text{Prob}_\ell = \frac{\exp\left[\varrho\mathcal{O}(R(\ell); I - k\ell)\right]}{\sum_{\ell=0}^{N-1} \exp\left[\varrho\mathcal{O}(R(\ell_n); I - k\ell_n)\right]}, \quad i = 0, 1, \ldots, N-1. \tag{49}$$

Taste heterogeneity is characterized concisely by parameter ϱ; the larger this parameter, the smaller the dispersion. With extreme taste heterogeneity, $\varrho = 0$, variance is infinite and choice is totally random and independent of utility evaluations: $\text{Prob}_{\ell_n} = 1/N$. With zero taste heterogeneity, $\varrho \to \infty$, variance is zero, and only the measured component of utility determines choice. That is, if $\mathcal{O}(R(\ell_i); I - k\ell_n) > \max_{\ell \neq \ell_n} \mathcal{O}(R(\ell); I - k\ell)$, then $\text{Prob}_{\ell_n} \to 1$.

Returning to the general model, let all individuals have identical incomes and each site have unit land area. Equilibrium may be defined in terms of a land rental function, $R(\ell)$, and a boundary $N_e < N-1$, such that

$$R^e(\ell_{N_e}) = R_a, \;\; R^e(\ell_n) \geq \; R_a, n < N_e, \tag{50}$$

where R_a denotes the exogenous rental land may earn in agricultural use at the city boundary, and the expected demand for land (housing) in every site is equal to supply,

$$h^*_\ell |\mathcal{I}| \text{Prob}_\ell = 1, \;\; \ell = 0, \ldots, N_e, \tag{51}$$

where the probability $P(\ell)$ that an individual would choose ℓ is given in (49). Mechanically, Equations (51) must be solved for all land rentals. Also, with the stochastic location model is associated an optimum expected utility, given by

$$\overline{\varpi} = \varrho^{-1}\ell_n \left[\exp \left[\sum_{\ell=0}^{N-1} \varrho \mathcal{O}(R(\ell); I - k\ell) \right] \right]. \tag{52}$$

The impact of heterogeneity may be examined in more detail if we specify the utility function. E.g., if the utility function is logarithmic, $\Omega(h,c) \equiv (1-\beta)\ell nc + \beta \ell nh$, then it turns out [*ibid.*, p. 330] that the equilibrium land rental rates obey:

$$\frac{R^e(\ell_i)}{R^e(\ell_{i+1})} = \left(\frac{I - k\ell_{i-1}}{I - k\ell_i} \right)^{\frac{\varrho+1}{\varrho\beta+1}}. \tag{53}$$

As ϱ increases, uncertainty decreases, that is preferences become more homogeneous and the exponent of the RHS increases, implying that the land rent gradient falls more sharply with distance. In the case of $\varrho \to \infty$, the exponent in the RHS of (53) tends to $\frac{1}{\beta}$, and we are in the Alonso–Mills–Muth case. In that case, utility is equalized across all locations. As the variance of utility shocks decreases, preferences become more homogeneous, rents increase within a certain distance from the CBD and decrease beyond that distance, thus making the urban structure more compact. One can think of preference heterogeneity as introducing correlated effects. The more similar individuals are, the closer together they would locate. This serves as a useful benchmark.

4.5 Income and preference heterogeneity

We assume that preferences are deterministic but individuals differ in terms of both incomes and taste across agents and imbed the problem in the simplified spatial setting of Anas (1990), presented in Section 4.4.1. Agents are thus described by means a joint density function of income and the preference parameter, $f(I, \beta)$. Individuals locate in concentric rings, that are defined by their distance from the CBD, $\ell = 0, \ell_1, \ldots, \ell_N$, and have area equal to one each. Self-selection of agents across these concentric rings is similar to what is analyzed by Epple and Platt (1998). Basically, the outcome may be analyzed easily because the single-crossing property of Ellickson (1971) holds. Stratification of individuals would be stable if individuals have no incentives to move across rings. The resulting stratification is described by defining, in (I, β) space, the loci of characteristics for individuals who are indifferent between locating in two successive rings, that is at distances $\ell_n, \ell_{n+1}, n = 1, \ldots, N-1$, respectively. Let $I = \gamma_n(\beta) \equiv \gamma_n(R_n, R_{n+1}; \ell_n, \ell_{n+1}; \beta)$, be the nth stratification envelope, given rents and distances for two successive rings, $(n, n+1)$ and radii ℓ_n, ℓ_{n+1}. That is:

$$(I - k\ell_n)R_n^{-\beta} = (I - k\ell_{n+1})R_{n+1}^{-\beta}, n = 0, \ldots, N_e.$$

Rewriting this condition yields:

$$\gamma_n(\beta) = k\frac{\ell_{n+1}\left(\frac{R_n}{R_{n+1}}\right)^{\beta} - \ell_n}{\left(\frac{R_n}{R_{n+1}}\right)^{\beta} - 1}. \tag{54}$$

It is straightforward to establish, by differentiating $\gamma_n(\beta)$ from (54) with respect to β, that the stratification envelope, $I = \gamma^n(\beta)$, is decreasing in both preference parameter β and the rent ratio R_n/R_{n+1}. Since the second derivative of the stratification envelope function with respect to the preference parameter is positive, then greater importance of housing in preference implies that the stratification loci are flatter and therefore all neighborhood become more homogeneous with respect to housing, *cet. par.*.

Since $R_N = R_{N_e} = R_a$, which is given exogenously as the opportunity cost of land, it follows that R_n declines with ℓ_n. For equilibrium in each ring, demand must equal supply. So, if the outermost ring is the Nth one, then

$$\int d\beta \int_{\gamma_{N-1}(\beta)}^{\infty} \frac{\beta(I - k\ell_N)}{R_a} f(I, \beta) dI = 1. \tag{55}$$

The LHS of this equation is monotonically increasing in and has a unique solution for $\frac{R_{N-1}}{R_N}$ the only unknown in (55). The solution may be written as: $\frac{R_{N-1}}{R_N} = \mathcal{R}(\ell_{N-1}, \ell_N)$. Similarly, housing market equilibrium in ring n requires:

$$\int d\beta \int_{\gamma_{n-1}(\beta)}^{\gamma_n(\beta)} \frac{\beta(I - k\ell_n)}{R_n} f(I, \beta) dI = 1, n = 1, \ldots, N-1. \tag{56}$$

The respective solution may be written as $\frac{R_{n-1}}{R_n} = \mathcal{R}\left(\frac{R_n}{R_{n+1}}; \ell_{n-1}, \ldots, \ell_N\right)$. Finally, the equation for equilibrium in ring 1 satisfies:

$$\int d\beta \int_{\min_{I_i, i \in \mathcal{I}}}^{\gamma_1(\beta)} \frac{\beta(I - k\ell_n)}{R_1} f(I, \beta) dI = 1. \tag{57}$$

The LHS in the above is monotonically decreasing in R_1, and therefore (57) determines $\frac{R_1}{R_2}$. Working recursively, we determine the entire sequence of housing rents and the number of rings, which are inhabited at equilibrium.

Finally, the stratification envelope for each of the rings is determined. Therefore, land use in this simple economy imply imperfect stratification: land use is mixed. The case of homogeneous preferences is, of course, a special case of this model. Once the stratification envelopes are determined, then the income distributions in each concentric ring follows as truncated marginal distributions. That is, the distribution of income in ring n, defined by radii ℓ_n, ℓ_{n+1}, is given by:

$$F_n(I) = \int d\beta \int_{\gamma_{n-1}(\beta)}^{\gamma_n(\beta)} f(I, \beta) dI.$$

Roughly speaking, incomes and housing preferences are negatively correlated within each neighborhood.

5. HIERARCHICAL MODELS OF LOCATION WITH SOCIAL INTERACTIONS

Next we discuss models of neighborhood choice, where the choice is over communities, defined as governmental jurisdictions that provide local public goods and finance them out of local taxes. This class of papers involve preferences over characteristics of other residents either directly or indirectly, via the package of public services. An important methodological difference characterizing these papers is a formal structure of ordered, or *hierarchical*, choice models. The analytics of these models in effect rest on a single unobservable, a preference parameter, that underlies individuals' choices and their endogenous separation into distinct communities. These models also employ data for populations aggregated at the governmental jurisdiction or community-level. In contrast, the class of papers examined earlier in the chapter typically employ micro data.

The analytical device that we saw is essential to understanding segregation in canonical urban models with a CBD, the single-crossing property, is a general tool of analysis in circumstances where agents self select. It continues to be useful here, as we see shortly.

A number of noteworthy papers, including several papers by Dennis Epple and coauthors such as Epple, Filimon and Romer (1984) and Epple and Romer (1990), and Benabou (1993; 1996a; 1996b), Durlauf (1996a; 1996b)), Epple and Platt (1998), and Epple and Romano (1998) emphasize the role of prices in bringing about rich sets of outcomes in the form of segregated or uniform equilibria. Becker and Murphy (2000) discuss important conceptual issues associated with the distinct roles of price rationing, zoning or segregation via government fiat in neighborhood choice, when neighborhoods differ in terms of exogenous amenities and endogenous socioeconomic composition.

Benabou (1996a) shows that with two communities that are equal in size, which are populated by individuals who differ in terms of a scalar characteristic and who value the community-specific average of the characteristic, the only stable equilibrium neighborhood outcome is segregated, if the marginal effect of the neighborhood contextual effect on the neighborhood price is increasing in the own characteristic. That is, for segregation it is required that individuals' willingness to pay for the neighborhood average of the individual characteristic be increasing with that individual characteristic. This is in effect a complementarity condition between the individual characteristic and its community-level counterpart. Durlauf (2004) notes that these models differ considerably in the approach to neighborhood structure, including fixed or endogenous size, the nature of housing market and how neighborhood membership is decided. For example, Durlauf (1996a) allows for endogenous number and size of neighborhoods, but imposes income requirements for neighborhood membership; Durlauf (1996b) shows that house prices may support a stratified outcome, when education is financed by a community-based income tax and is sensitive to the neighborhood income distribution.

5.1 The Epple-Sieg class of models

This class of papers include Epple and Sieg (1999), Epple, Romer and Sieg (2001), and Calabrese *et al.* (2006). These authors start from a given total population, defined in terms of the distribution of individuals' demographic characteristics, who allocate themselves via the housing market to a given number of distinct communities. The important difference that distinguishes these models is provision of a community-level public good, which is financed through local taxation of housing. Covariation of preferences and income induces sorting. In the Manski typology, such sorting is induced by correlated effects. Individuals who are similar in terms of preferences and incomes tend to sort in the same community, with housing prices reflecting all available information and helping ration access to communities. However, sorting is imperfect in these models, just as in real economic life: individuals with the same income may be found in different communities.

We assume that all individuals make decisions at the same time about where to locate among $\ell = 1, \ldots, \mathcal{L}$ communities. Individuals' preferences are defined in terms of their indirect utility functions, as functions of individual income I, of the price of housing in community ℓ, P_ℓ, an observable variable in principle, of the tax price for a local public good, g_ℓ, and $\varepsilon > 0$, an individual characteristic. Following Epple and Sieg (1999), we assume an indirect utility function for a household with income I residing in neighborhood ℓ, of the form:

$$\Omega(I, P_\ell, g_\ell; \varepsilon) \equiv \left[\varepsilon g_\ell^{\psi} + \left[\exp\left(\frac{I^{1-\delta} - 1}{1-\delta} \right) \exp\left(-B \frac{P_\ell^{\mu+1} - 1}{1+\mu} \right) \right]^{\psi} \right]^{\frac{1}{\psi}}, \tag{58}$$

where: $\psi < 0$, $\mu < 0$, $\delta > 0$, and $B > 0$ are parameters that are constant across all households.

To see how the assumption about preferences according to (58) serves to sustain sorting across neighborhoods, we may think of "indirect indifference" curves in (P_ℓ, g_ℓ) space. Their slopes given by:

$$\frac{\partial g}{\partial P}\bigg|_{\Omega=\text{const}} = \frac{\left[\exp\left(\frac{I^{1-\delta}-1}{1-\delta}\right)\exp\left(-B\frac{P_k^{\mu+1}-1}{1+\mu}\right)\right]^{\psi} Bp^{\mu}}{\varepsilon g^{\psi-1}} > 0. \tag{59}$$

These indifference curves are essential in characterizing neighborhood sorting for the following reason. Since they are monotonic in I and ε, they satisfy the single-crossing property with respect to income, I, and to the taste parameter, ε, given I. As Epple and Sieg, *op. cit.*, show, this property is crucial for obtaining separating equilibria, with respect to both income, I and the taste parameter, ε. To see this intuitively, consider the indirect indifference curves for two values of income, I', I'', $I' < I''$, with the same value of ε. As the indifference curve for I'' cuts the one for I' from below, individuals with incomes equal to I'' are willing to bid a higher value to locate in a community with higher value of P_ℓ, holding g_ℓ constant, would be populated by households with higher incomes.

We index the $\mathcal{L}$ neighborhoods in individuals' opportunity sets, so that: $g_\ell < g_{\ell+1}, P_\ell < P_{\ell+1}, \ell = 0, \ldots, \mathcal{L}-1$. According to Epple and Sieg, *op. cit.* p. 651, there must be an ordering of communities that must be confirmed at equilibrium.[21] We assume that this indexing coincides with the equilibrium ordering. We work out the specifics of selection, which is likely to emerge under preferences (58). Next we seek to characterize the marginal density function for income in community $\ell, f_\ell(I)$ in terms of a given joint density function of preferences characteristics and income, $f(I, \varepsilon)$, across the population of the entire metropolitan area.

The set of individuals $j \in \mathcal{I}_\ell$ who reside in community ℓ are characterized by the set of values (I_j, ε_j) such that:

$$V(I_j, P_{\ell-1}; g_{\ell-1}; \varepsilon_j) < V(I_j, P_\ell; g_\ell; \varepsilon_j) \leq V(I_j, P_{\ell+1}; g_{\ell+1}; \varepsilon_j). \tag{60}$$

We follow Epple and Sieg (1999) but simplify by setting $\delta = 1$ in which case $\lim_{\delta=1} : \left(\frac{I^{1-\delta}-1}{1-\delta}\right) = \ell nI$. It turns out that the boundary of communities ℓ and $\ell+1$ in $(\ell nI, \ell n\varepsilon)$ space is the straight line given by $\ell n\varepsilon - \psi \ell nI = C_k$. Conditions (60) are transformed into:

$$C_{\ell-1} + \psi \ell nI < \ell n\varepsilon \leq C_\ell + \psi \ell nI, \tag{61}$$

[21] This ordering must satisfy *boundary indifference, income stratification, and ascending bundles*. That is, if $P_i > P_j$, then $g_i > g_j$, iff community i is populated by higher income people than community j.

where the C_k's are auxiliary variables defined by

$$C_\ell \equiv \ell n \left(\frac{\exp\left[\frac{-\psi}{\mu+1}\left(BP_{\ell+1}^{\mu+1}-1\right)\right] - \exp\left[\frac{-\psi}{\mu+1}\left(BP_{\ell}^{\mu+1}-1\right)\right]}{g_\ell^{\psi} - g_{\ell+1}^{\psi}} \right), \ell = 1, \ldots, \mathcal{L}-1. \quad (62)$$

Note that C_ℓ is increasing in g_ℓ and P_ℓ. Our assumptions about the ranking of the $\mathcal{L}$ neighborhoods imply that the C_ℓ's, which are functions of prices and public good levels, $(g_\ell, P_\ell; g_{\ell+1}, P_{\ell+1})$, and of parameters, satisfy $C_{\ell+1} > C_\ell$. For completeness, we define $C_0 = -\infty$, and $C_{\mathcal{L}} = \infty$. Therefore, all information that is relevant for sorting of individuals into communities is encapsulated in the auxiliary variables $C_\ell, C_\ell = 1, \ldots, \mathcal{L}$.

The number of people who reside in community ℓ, as a percentage of the total population of the metropolitan area, is given by the probability over the set defined by (61), which is a function of $(C_{\ell-1}, C_\ell)$, and of the parameters of the joint distribution of income and taste parameter, $f(I, \varepsilon)$.

The Epple and Sieg estimation method simply matches community-specific populations with those predicted from the sorting model. The package of public goods and the associated tax rates may be endogenized once community decision making is modeled. This is accomplished, for example, by Epple, Romer and Sieg (2001), who also employ the estimation strategy of Epple and Sieg (1999) again for the communities of the Boston metropolitan area. Specifically, they use community-level data on population size, number of households, mean income, median income, education expenditure, property tax rate, median property value, median gross rent and fraction of renters. They estimate all behavioral parameters and the distribution of income jointly with ε, the behavioral parameter that indexes heterogeneity. The estimation approach rests on an index for a community-level public good (services), in their case spending on education and community crime rates, and an instrument for a community-specific housing price that may be estimated from data on housing expenditures (in their approach measured by owner-occupied housing or rents), incomes and property tax rates. Community-specific hedonic price equations could also deliver prices that may be used in the estimation of the community-sorting model.

The properties of the Epple-Sieg sorting model, as demonstrated by (61), are summarized as follows. First, given the value of the taste parameter ε, individuals are perfectly stratified across communities in terms of income; and second, given income, individuals are perfectly stratified across communities in terms of taste. See Epple and Sieg, *op. cit.*, p. 653, Figure 1. Consequently, individuals' taste and income are *positively correlated* across communities, that is as we move along a 45°–degree line away from the origin. Individuals' taste for the community attribute and income are *negatively correlated* within communities, that is as we move within each of the strips that define communities. See Epple and Platt (1998), p. 31, Figure 3. Higher income people must have lower taste for the community attribute than lower income people who reside in the same community. This is an interesting feature of the theoretical model that has bearing upon the estimation.

Given parametric assumptions on the joint distribution of income and tastes for the population of the metropolitan area, the model determines a joint distribution of income and taste parameters for every community. If the model is evaluated at the correct parameter values, the difference between the empirical quantiles of the income distributions observed in the data and the quantiles predicted by the model should be small. This provides the rationale for the first stage of the estimation. Heterogeneity in tastes and income in the metropolitan population, together with self-selection of households into municipalities, means that income distributions will differ across municipalities in the metropolitan area. This allows them to estimate the parameters of the income distribution, the correlation of income with the taste parameter, the ratio of $\psi/\sigma_{\ell n\varepsilon}$, and the income elasticity of housing demand. The remaining parameters of the model are estimated by matching the observed distribution of community-level tax rates, expenditures, and imputed rents. The baseline model with no peer effects fits reasonably well, but overstates expenditure in the lower-income communities and understates them in the higher-income communities. The model over-predicts rents in the lower-income communities and under-predicts them in the higher-income communities.

Epple and Sieg find that 89% of the total variance of income in the Boston metropolitan area is accounted for by within community variance.[22] They interpret this as evidence of substantial unobserved heterogeneity in preferences for local public goods, in their case a linear combination of school quality and crime. Their estimation of the joint distribution of the taste parameter and income is done with great precision. All of the parameters of their model are identified, although the elasticity of housing demand is identified "from nuances of the functional form" rather than from information on housing expenditures. Their estimates of μ are in the range of -0.30 to -0.50.

Epple, Romer and Sieg (2001) test the political economy part of the model, that is, whether observed levels of public expenditures satisfy necessary conditions implied by majority rule in a general equilibrium model of residential choice. Again, the model determines a joint distribution of income and taste parameters for every community. The first stage of the estimation strategy is based on the idea that the difference between the quantiles of the income distributions, as observed in the data, and the quantiles, as predicted by the model, should be small if the model is evaluated at the correct parameter values. Their approach treats the auxiliary variables, the C_ℓ's, as unknown parameters and estimates the model using a minimum distance estimator. The basic idea of their locational equilibrium estimator is to match the levels of public

[22] A similar observation was made by Hardman and Ioannides (2004) and Ioannides (2004).

good provision implied by the first stage estimates with those observed in the data, conditional on differences in housing prices. The estimator controls for observed and unobserved heterogeneity among households, observed and unobserved characteristics of communities, and the potential endogeneity of housing prices and public good expenditures. They estimate the structural parameters of the model using 1980 data from the Boston MSA. A key contribution here is to show that it is in fact possible to estimate consistently the underlying parameters using orthogonality conditions derived from the assumption of majority rule. They extend the analysis to estimate jointly the orthogonality conditions for majority rule and the orthogonality conditions for locational equilibrium. The first tests a myopic voting model, according to which voters ignore all effects of migration and treat the population boundaries of the communities as fixed. In other words, they believe that the distribution of households across communities is not affected by a change in public good provision. This model maps into the assumptions that the net-of-tax price of housing, community population, and the aggregate housing demand are fixed. This model is rejected. More sophisticated voting models based on utility taking provide a potential explanation of the main empirical regularities. This variation of their model incorporates mobility into the computation of the government possibility frontier. Their estimates suggest that the implied tradeoff between public good and housing price are similar for the mid range of the communities distribution and yet sufficiently different at the both ends to support the conclusion of a much better fit for the utility-taking model.

5.2 Peer effects

The nature of the attractiveness of different communities via preference over a publicly provided community-level good explains sorting due to correlated effects, that is, individuals with similar preferences will seek to reside in the same communities. It falls short, however, of accounting for neighborhood effects that would reflect the characteristics of individuals who self-select into communities. Calabrese, Epple, Romer and Sieg (2006) specify the community-specific attribute in (58) as a neighborhood effect. The quality of the public good is defined as its physical quantity, $\bar{g}_\ell$, adjusted by the mean income of community ℓ, $\bar{I}_\ell$, relative to that of the entire economy, $\bar{I}$. That is:

$$g_\ell \equiv \bar{g}_\ell \left(\frac{\bar{I}_\ell}{\bar{I}} \right)^{\phi}, \tag{63}$$

where $\phi \geq 0$, is a parameter. In each community, the cost of the public good is financed by a proportional tax on housing expenditure, which is decided by voting, as before. Although the authors refer to this feature of their model as a peer effect, it is actually a contextual effect according to the Manski typology. See also Luttmer (2005) for individuals' preferences over relative earnings.

These authors' empirical approach allows them to impose all restrictions that arise from locational equilibrium models with myopic voting simultaneously on the data generating process. Recall the specific features of the baseline model discussed above, namely that it tends to over-predict housing values and under-predict expenditures in lower-income communities, and the other way around for higher-income ones. These cause a severe under-prediction of tax rates in poorer communities and over-prediction in higher-income ones. Interestingly, these problems are ameliorated when peer effects are included. Expressing the exponent of the quality function ϕ in relative terms, their estimates imply that the peer effects are 2.5 times as important as spending. Introducing peer effects into the model specification markedly improves the fit of the model. The model with peer effects not only explains expenditures, but also tax rates and tax bases (rents) reasonably well. In fact, these authors find that the correlation between actual and predicted tax rates is 0.747, instead of -0.67 in the baseline model. The extended myopic voting model that allows for peer effects in public good provision yields a good fit to the data.

Overall, they find that their relatively simple model of sorting and public good provision fits the data on community income distributions, housing expenditures, public good provision levels, and property tax rates reasonably well and that peer effects may be important components in determining the quality of local public good provision.

Recall that Epple *et al.* (2001), working with a model where the public good is determined through voting, found that the parameter estimates from the locational equilibrium and voting equilibrium components of the model led to different results that were difficult to reconcile. The results of Calabrese *et al. op. cit.*, eliminates the apparent inconsistency. Clearly, peer effects enhance the flexibility of the model. Lower income communities may have a lower tax base but also face a lower quality of the public good on account of lower relative income.

The authors discuss whether there exist any explanations other than peer effects that might explain their earlier results. One possibility is features of the state fiscal system that introduce a wedge between a community's property tax as the marginal source of funds for increasing educational expenditures in the model and the actual one. However, they argue that that was not the case in Massachusetts at the time of their data, when state aid was based on local property tax base and school enrollments had a redistributive effect.

The authors also discuss how peer effects would operate, if they are present. Peer effects may be due either to production or consumption externalities. However, they could proxy for endogenous social effects, in which case the model to be estimated may be considered a reduced form in the context of the social interactions literature. Peer effects conceived at the community level that have received particular emphasis are those operating through schools. If that is the primary mechanism for peer effects, then one would expect that parents of school-age children would tend to locate where the peer variable is higher, and households without children would locate where the peer variable is relatively low. Mean household income has a correlation coefficient 0.57 with the fraction of households that are families, 0.32 with school enrollment per household, and −0.38 with the fraction of the population over 65 years of age. To the extent that peer effects operate through education of children, the education of the adult population may provide a measure both of the value attached to education and of the resources available to facilitate student learning. The authors investigate this by regressing, across communities, the logarithm of mean household income on a constant, the fraction of the population with high school education, and the fraction with more than high school education. All estimated coefficients are highly significant, and the $R^2 = 0.83$. Thus, their peer variable, mean community income, is strongly related to the education of the community population. While the evidence is only suggestive, it is consistent with the peer effects interpretation of the results from their structural estimation.

As Epple, Peress, and Sieg (2010) emphasize that, in this class of community-choice models, the hierarchical nature of community choice is an important ingredient of the model. The ascending bundles property of public goods and tax rates determines sorting by jointly summarizing the value for households of living in different communities and rationing demand accordingly. In this paper, the authors discuss identification and estimation of hierarchical equilibrium models in semi-parametric frameworks and extend into a broad class of nonhierarchical models. Their results suggest that identification of nonhierarchical models may ultimately have to rely on stronger assumptions on the distribution of unobserved tastes for public goods than those used in hierarchical models.

These authors apply the Epple *et al.* type of model to community choice and housing demand, using data on 93,763 properties in 150 or so communities (which include the wards of the City of Pittsburgh) in Allegheny County, Pennsylvania. The data include detailed price and quality characteristics for dwelling units. In addition, they use data on community characteristics based on educational standards, crime and travel time to the city center. They estimate housing prices, using dwelling unit data, and the parameters of housing demand, using Census data aggregated at the community level, taking advantage of the quantiles of housing consumption and income distributions available within the Census data. Next, they estimate a function of product quality, as a function of housing prices (imputed rents). They characterize the observed sorting of household types across communities, implied by these probabilities, by plotting the share of different types of households with given incomes who live in communities which have housing prices less than or equal to the price of community ℓ. These plots show that households with children are more responsive to differences in housing prices (and local public good provision) than those without children, households with and without children being the two specific types used to demonstrate the method. The sorting of households with children exhibits more stratification by income than those without children. They also find that households with children and income levels below the mean metropolitan income are more likely to live in cheaper communities than households without children. The opposite is true for households with high levels of incomes. High-income households with children have stronger preferences for high price (and high amenity) communities than households without children. They complete the identification of the model and thus demonstrate that their framework is rich enough to uncover important features of residential sorting which would have obscured by conventional parametric estimators.

We have dwelled extensively on the Epple-Sieg approach because we regard it as an important framework in evaluating how attractiveness of communities may be related to social interactions within communities, even when that might not be specifically spelled out and controlled for. When considered against hedonic-type models, the housing price again does double duty, by pricing housing and the right to live in a community. In the case of the earlier Epple and Sieg approach, the discreteness of the choice set limits the identification of the model, whereas in the hedonic approach, and the latest Epple and Sieg type application, just discussed, the first order conditions are exploited in order to estimate the hedonic model. We return to this further below. We note, in particular, that sorting works here only through individuals' valuation of community-specific amenities and not neighborhood effects as such. Of course, when applied to smaller communities, those two approaches may be close substitutes.

Finally, we discuss briefly an approach that combines micro with aggregate data [Ioannides and Schmidheiny (2006)]. These authors develop a model of community choice as discrete choice by combining features of the approach by Berry, Levinsohn and Pakes (1995; 2004) to the choice of differentiated goods with the approach of Epple and Sieg (1999) and of Epple *et al.* (2001; 2010). The paper reports estimation results that involve an iterative procedure consisting of two stages. At a first stage, they use information on the *joint* distribution of individual household characteristics in a metropolitan area, which in their case are obtained from the American Housing Survey Boston metropolitan area micro sample, to predict population shares and moments of household characteristics and to match them with observed population shares and moments by community by means of a generalized method of moments method. This stage uses estimates of effects reflecting interactions between individual and community characteristics and estimates community-specific intercepts. These intercepts serve as sufficient statistics for the estimation, at a second stage, of coefficients expressing the effects of community-specific characteristics by regressing them on the *marginal* distributions of household characteristics of the different communities in Boston metropolitan area in 1980. The latter is the same data set from the US Census that Epple *et al.* have also used, but is augmented by means of community-specific housing prices, which are obtained from the record of all transactions in the respective communities within the Boston metropolitan area. Their results are noteworthy because they demonstrate that use of two public sources of data may circumvent the need of confidential data that some other research has relied on. It also demonstrates the potential for use of additional information in the form of additional moment conditions associated with the distributions of observable household characteristics.

6. CONCLUSION

This chapter presents a number of approaches that economists have used in studying neighborhood effects in housing markets. It also aims, to an extent that might be possible, at unifying these approaches. It emphasizes how individuals in effect choose neighborhood effects, or more generally social interactions, by means of their location decisions. Instead of emphasizing the role of neighborhood effects as mere externalities that might interfere with locational equilibrium, this chapter looks constructively at neighborhood effects and focuses on what we have learned empirically about their role by observing locational decisions or patterns along with individual and group characteristics.

We examine several classes of models, which economists have relied upon in exploring the role and empirical significance of neighborhood effects that originate in housing markets. So, we take the concept of neighborhood effects quite literally as arising in residential neighborhoods. For precisely this reason, the chapter emphasizes how we may detect empirically the presence of neighborhood effects when they may be priced by housing markets and be capitalized into housing values and rents. The chapter thus focuses on models of primarily empirical relevance that may help identify neighborhood effects and discusses actual empirical findings.

The first class of models we examine involves models of choice over discrete sets of individual dwelling units that allow for a multidimensional bundle of characteristics. Models of these types, which borrow from the industrial organization literature and especially the Berry-Levisohn-Pakes characteristics-based models, have been developed by Bayer and a number of coauthors. They lead naturally to hedonic models while sorting is accounted for. In the context of the neighborhood effects literature, these models allow for endogenous contextual effects. We pursue further a conceptually related approach, due to Nesheim, that also endogenizes contextual effects but involves much lower dimensionality. That is, individuals choose neighborhoods while recognizing that their neighbors' characteristics along with their own determine educational outcomes for their children. This approach allows us to obtain equilibrium housing price functions that are consistent with hedonic valuation of neighborhood attributes.

The paper turns next to aspatial models of neighborhood choice, with endogenous and contextual neighborhood effects, and housing demand (with housing is measured as a scalar) as joint decisions. It emphasizes a model, due to Ioannides and Zabel, which is designed to utilize individual and neighborhood-level data at several levels of aggregation. This approach links naturally with hierarchical choice models. Hierarchical models of neighborhood location in the presence of social interactions, as developed by Epple and co-authors, describe communities in terms of a low-dimensional vector of attributes that may be aggregated into a public good. Individuals' choice of community is subject to community-specific housing price and tax rate, which at equilibrium must sustain individuals' choices. This approach is designed to utilize data that are aggregated at the community-level, along with information on the community-specific distributions of various sociodemographic characteristics of individuals. When considered against hedonic-type models, the housing price again does double duty, by pricing housing and rationing "admission" into a community. In these models, sorting across communities works either only through individuals' valuation of community-specific amenities or with community-specific amenities combined with neighborhood effects in the form of peer effects.

The chapter takes up neighborhood effects within the canonical Alonso–Mills–Muth urban model with a well-defined spatial structure and individuals commuting to a predetermined central business district. If individuals differ with respect to income or preference characteristics, then the standard urban model implies segregation. The paper discusses extensions of the model that allow for amenities, that may be exogenous or endogenous and are spatially dispersed, and naturally influence individuals' location decisions and the associated housing price structure.

Although many of the models presented in this chapter may seem very special, they provide, we think, the building blocks for a full understanding of neighborhood effects in housing markets. And quite importantly, a full understanding of neighborhood effects in housing markets is crucial for evaluating the allocative role of prices when they ration admission to communities and neighborhoods. This natural interface with hedonic theory, which has received a lot of attention here, is essential in assessing what can be learned from empirical analyses.

At the same time, many of the results discussed here suggest that the concept of neighborhood is multifaceted. Individuals indeed care about the quality of neighborhoods at several levels (scales). The information at the such traditional levels as clusters, tracts and MSAs can be highly correlated, but there is also independent information at those different levels. This accords with the notion that different small neighborhoods have different characters, and their proximity to one another confer character to higher-level neighborhoods that contain them. The literature has yet to develop effective measures of these aspects; but hopefully, the approaches discussed in this chapter constitute a start.

REFERENCES

Alonso, W., 1964. Location and Land Use: Toward a General Theory of Land Rent. Harvard University Press, Cambridge, MA.

Anas, A., 1990. Taste Heterogeneity and Urban Spatial Structure: the Logit Model and Monocentric Theory Reconciled. J. Urban Econ. 28 (3), 318–335.

Anas, A., Feng, C.M., 1988. Invariance of Expected Utilities in Logit Models. Econ. Lett. 27, 41–45.
Anderson, S.P., de Palma, A., Thisse, J.-F., 1992. Discrete Choice Theory of Product Differentiation. MIT Press, Cambridge, MA.
Angrist, J.D., Pischke, J.S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton, NJ.
Bayer, P., Ross, S.L., 2009. Identifying Individual and Group Effects in the Presence of Sorting: A Neighborhood Effects Application. NBER working paper No. 12111.
Bayer, P., McMillan, R., Rueben, K., 2005. An Equilibrium Model of Sorting in an Urban Housing Market. NBER working paper No. 10865.
Bayer, P., Timmins, C., 2005. On the Equilibrium Properties of Locational Sorting Models. J. Urban Econ. 57 (3), 462–477.
Bayer, P., Timmins, C., 2007. Estimating Equilibrium Models of Sorting Across Locations. Economic Journal 117, 353–374.
Bayer, P., McMillan, R., Rueben, K., 2004a. Residential Segregation in General Equilibrium. working paper, Yale University, March.
Bayer, P., McMillan, R., Rueben, K., 2004b. What Drives Racial Segregation? New Evidence Using Census Microdata. J. Urban Econ. 56, 514–535.
Bayer, P., Ferreira, F.V., McMillan, R., 2007a. A Unified Framework for Measuring Preferences for Schools and Neighborhoods. J. Polit. Econ. 115 (4), 588–638.
Bayer, P., Ferreira, F.V., McMillan, R., 2007b. Tiebout Sorting, Social Multipliers and the Demand for School Quality. NBER working paper No. 10871.
Bayer, P., Ross, S.L., Topa, G., 2008. Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. J. Polit. Econ. 116 (6), 1150–1196.
Bayer, P., McMillan, R., Rueben, K., 2009. An Equilibrium Model of Sorting in an Urban Housing Market. NBER working paper No. 10865.
Becker, G.S., Murphy, K.M., 2000. Segregation and Integration in Neighborhoods. In: Becker, G.S., Murphy, K.M. (Eds.), Social Economics. Harvard University Press, Cambridge, MA.
Benabou, R., 1993. Workings of a City: Location, Education, and Production. Q. J. Econ. 108 (3), 619–652.
Benabou, R., 1996a. Equity and Efficiency in Human Capital Investment: The Local Connection. Rev. Econ. Stud. 63 (2), 237–264.
Benabou, R., 1996b. Heterogeneity, Classification, and Growth: Macroeconomic Implications of Community Structure and School Finance. Am. Econ. Rev. 86 (3), 584–609.
Berry, S.T., 1994. Estimating Discrete-Choice Models of Product Differentiation. RAND J. Econ. 25 (2), 242–262.
Berry, S.T., Levinsohn, J., Pakes, A., 1995. Automobile Prices in Market Equilibrium. Econometrica 63 (4), 841–890.
Berry, S.T., Levinsohn, J., Pakes, A., 2004. Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market. J. Polit. Econ. 112 (1), 68–105.
Bierlaire, M., Bolduc, D., McFadden, D., 2003. Characteristics of Generalized Extreme Value Distributions. Technical report, Econometrics Laboratory, University of California, Berkeley.
Bisin, A., Moro, A., Topa, G., 2009. The Empirical Content of Models with Multiple Equilibria in Economies with Social Interactions. New York University, mimeo, September.
Blackley, P., Ondrich, J., 1988. A Limiting Joint-Choice Model with Discrete and Continuous Housing Characteristics. Rev. Econ. Stat. 70, 266–274.
Black, S., 1999. Do Better Schools Matter? Parental Valuation of Elementary Education. Q. J. Econ. 114 (2), 577–599.
Blume, L.E., Brock, W.A., Durlauf, S.N., Ioannides, Y.M., 2010. Identification of Social Interactions. In: This Handbook.
Brock, W.A., Durlauf, S.N., 2001. Interactions-Based Models. In: Heckman, J.J., Leamer, E. (Eds.), Handbook of Econometrics. 5, North-Holland, Amsterdam.
Brock, W.A., Durlauf, S.N., 2002. A Multinomial Choice Model of Neighborhood Effects. Am. Econ. Rev. 92, 298–303.

Brock, W.A., Durlauf, S.N., 2007. Identification of Binary Choice Models with Social Interactions. J. Econom. 140, 52–75.

Brooks-Gunn, J., Duncan, G., Klebanov, P., Sealand, N., 1993. Do Neighborhoods Affect Child and Adolescent Development? Am. J. Sociol. 99 (2), 353–395.

Brueckner, J.K., Thisse, J.-F., Zenou, Y., 1999. Why Is Central Paris Rich and Downtown Detroit Poor? An Amenity-Based Theory. Eur. Econ. Rev. 43, 91–107.

Calabrese, S., Epple, D., Romer, T., Sieg, H., 2006. Local Public Good Provision: Voting, Peer Effects, and Mobility. J. Public Econ. 90, 959–981.

Dougherty, P.J., 2002. Who Is Afraid of Adam Smith? How The Market Got its Soul!. John Wiley and Sons, New York.

Dubin, J.A., McFadden, D.L., 1984. An Econometric Analysis of Residential Electrical Appliance Holdings and Consumption. Econometrica 52, 345–362.

Durlauf, S.N., 1996a. A Theory of Persistent Income Inequality. J. Econ. Growth 1 (1), 75–93.

Durlauf, S.N., 1996b. Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality. In: Barnett, W. *et al.,* (Ed.), Dynamic Disequilibrium Modelling. Cambridge University Press, Cambridge.

Durlauf, S.N., 2004. Neighborhood Effects. In: Henderson, J.V., Thisse, J.-F. (Eds.), Handbook of Regional and Urban Economics. 4, Cities and Geography, North-Holland, Amsterdam.

Durlauf, S.N., Ioannides, Y.M., 2010. Social Interactions. Ann. Rev. of Econ. September.

Ekeland, I., Heckman, J.J., Nesheim, L., 2004. Identification and Estimation of Hedonic Models. J. Polit. Econ. 112 (1), S60–S109.

Ellickson, B., 1971. Jurisdictional Fragmentation and Residential Choice. American Economic Review, Papers and Proceedings 61 (2), 334–339.

Ellickson, B., 1981. An Alternative Test of the Hedonic Theory of Housing Markets. J. Urban Econ. 9, 56–79.

Epple, D., Filimon, R., Romer, T., 1984. Equilibrium among Local Jurisdictions: Toward an Integrated Treatment of Voting and Residential Choices. J. Public Econ. 24 (3), 281–304.

Epple, D., Gordon, B., Sieg, H., 2010. A New Approach to Estimating the Production for Housing. Am. Econ. Rev. 100 (3), 905–924.

Epple, D., Peress, M., Sieg, H., 2010. Identification and Semi-Parametric Estimation of Equilibrium Models of Local Jurisdictions. Am. Econ. J. Microeconomics, forthcoming.

Epple, D., Platt, G.J., 1998. Equilibrium and Local Redistribution in an Urban Economy when Households Differ in Both Preferences and Incomes. J. Urban Econ. 43, 23–51.

Epple, D., Romano, R., 1998. Competition between Private and Public Schools and Peer Group Effects. Am. Econ. Rev. 88 (1), 33–62.

Epple, D., Romer, T., 1990. Mobility and Redistribution. J. Polit. Econ. 99 (4), 828–858.

Epple, D., Romer, T., Sieg, H., 2001. Interjurisdictional Sorting and Majority Rule: An Empirical Analysis. Econometrica 69 (6), 1437–1465.

Epple, D., Sieg, H., 1999. Estimating Equilibrium Models of Local Jurisdictions. J. Polit. Econ. 107, 645–681.

Fox, J.T., 2007. Semiparametric Estimation of Multinomial Discrete-Choice Models Using a Subset of Choices. RAND. J. Econ. 38 (4), 1002–1019.

Fujita, M., 1989. Urban Economic Theory: Land Use and City Size. Cambridge University Press, Cambridge.

Graham, B.S., 2009. Endogenous Neighborhood Selection, the Distribution of Income and the Identification of Neighborhood Effects. Department of Economics, New York University.

Hardman, A.M., Ioannides, Y.M., 2004. Neighbors' Income Distribution: Economic Segregation and Mixing in US. J. Hous. Econ. 13, 368–382.

Ioannides, Y.M., 2003. Empirical Nonlinearities and Neighborhood Effects in the Intergenerational Transmission of Human Capital. Appl. Econ. Lett. 10, 535–539.

Ioannides, Y.M., 2004. Neighborhood Income Distributions. J. Urban Econ. 56, 435–457.

Ioannides, Y.M., 2008. Full Solution of an Endogenous Sorting Model with Contextual and Income Effects. Working paper, Department of Economics, Tufts University, August.

Ioannides, Y.M., 2010. From Neighborhoods to Nations: The Economics of Social Interactions. In progress, February.
Ioannides, Y.M., Schmidheiny, K., 2006. Estimating Neighborhood Choice by Combining Individual and Neighborhood Data. Tufts University and Universitat Pompeu Fabra; presented at the North American Meeting of the Regional Science Association International, Toronto. November.
Ioannides, Y.M., Zabel, J.E., 2008. Interactions, Neighborhood Selection and Housing Demand. J. Urban Econ. 63, 229–252.
Jaïbi, M.R., ten Raa, T., 1998. An Asymptotic Foundation for Logit Models. Reg. Sci. Urban Econ. 28 (1), 75–90.
Kiel, K.A., Zabel, J.E., 2009. Location: The 3L Approach to House Price Determination. J. Hous. Econ. 17, 175–190.
Kremer, M., 1997. How Much Does Sorting Increase Inequality? Q. J. Econ. 112 (1), 115–139.
Lee, D.S., Lemieux, T., 2010. Regression Discontinuity Designs in Economics. J. Econ. Lit. forthcoming.
Luttmer, E.F.P., 2005. Neighbors as Negatives: Relative Earnings and Well-Being. Q. J. Econ. 120 (3), 963–1002.
Manski, C.F., 1993. Identification of Endogenous Social Effects: The Reflection Problem. Rev. Econ. Stud. 60, 531–542.
Manski, C.F., 2000. Economic Analysis of Social Interactions. J. Econ. Perspect. 14 (3), 115–136.
McFadden, D.F., 1978. Modelling the Choice of Residential Location. In: Karlqvist, A., Lundqvist, L., Snickars, F., Weibull, J. (Eds.), Spatial Interaction Theory and Planning Models. North-Holland, Amsterdam.
Mills, E.C., 1967. An Aggregative Model of Resource Allocation in a Metropolitan Area. Am. Econ. Rev. 57 (2), 197–210.
Muth, R.F., 1969. Cities and Housing: The Spatial Pattern of Urban Residential Land Use. The University of Chicago Press, Chicago.
Nesheim, L., 2002. Equilibrium Sorting of Heterogeneous Consumers across Locations: Theory and Empirical Implications. CeMMAP working paper CWP08/02, University College, London. March.
Ogawa, H., Fujita, M., 1980. Equilibrium Land Use Patterns in a Non-monocentric City. J. Reg. Sci. 20, 455–475.
Ogawa, H., Fujita, M., 1989. Nonmonocentric Urban Configurations in a Two-Dimensional Space. Environ. Plan. A 363–374.
Quigley, J.M., 1985. Consumer Choice of Dwelling, Neighborhood and Public Services. Reg. Sci. Urban Econ. 15, 41–63.
Rosen, S., 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. J. Polit. Econ. 82, 34–55.
Rosen, S., 2002. Markets and Diversity. Am. Econ. Rev. 92 (1), 1–15.
Samuelson, P.A., Swamy, S., 1974. Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis. Am. Econ. Rev. 64, 566–593.
Sieg, H., Smith, V.K., Banzhaf, H.S., Walsh, R., 2002. Interjurisdictional Housing Prices in Locational Equilibrium. J. Urban Econ. 52, 131–153.
Zabel, J.E., 2004. The Demand for Housing Services. J. Hous. Econ. 13 (1), 16–35.