Full Solution of an Endogenous Sorting Model with Contextual and Income Effects

Yannis M. Ioannides*

Department of Economics

Tufts University

Medford, MA 02155

Telephone: 617 627 3294 Fax: 617 627 3917

`Yannis.Ioannides@Tufts.edu` `http://www.tufts.edu/˜yioannid/`

August 17, 2008

Abstract

The paper solves analytically a problem where individuals sort themselves into neighborhoods because they value average schooling among adults in the neighborhood. The paper extends results by Nesheim (2002) but with the addition of income effects on neighborhood choice. Individuals value housing, non-housing consumption, and expected schooling of their children. The latter depends on parental schooling, on a child's ability, and on average schooling in the neighborhood. Neighborhood choice trades off non-housing consumption with children's expected schooling. Individuals choose neighborhoods recognizing that their neighbors' characteristics are correlated with their own. The equilibrium housing price is associated with endogenous sorting and also allows computation of neighborhood distributions of income, of schooling and other variables of interest.

* The results of this paper were reported originally in my "Neighborhood Effects and Housing," version March 3, 2008, *Handbook of Social Economics,* Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, eds., Elsevier, forthcoming. I thank Bryan Graham for sharing with me his "On the Identification of Neighborhood Externalities in the Presence of Endogenous Neighborhood Selection," June 26, 2008. The usual caveats apply. `IoannidesSolutionSorting-August-17-08.tex`

Contents

1 Introduction

The paper solves analytically a sorting problem, as originally formulated by Nesheim (2002) but modified to also allow for income effects in neighborhood choice. Individuals value housing and expected schooling of their children. The latter depends on parent's schooling, on a child's ability, and on average parental schooling in the neighborhood. Neighborhood choice trades off housing with advantageous contextual effects in the production of schooling. Individuals choose neighborhoods while recognizing that their neighbors' characteristics are correlated with their own. The equilibrium housing price function that we obtain is associated with endogenous sorting and also allow us to compute neighborhood distributions of income and of schooling.

The paper succeeds in going beyond Nesheim's analytical results without resorting to numerical simulation by slightly varying one of his assumptions. This in a sense "breaks" the symmetry of his assumptions but we regard doing so as inessential in view of the purpose of the paper and the clarity of the analytical results.

Endogenous sorting is at the heart of a number of problems that have attracted considerable interest recently. For one, self-selection into neighborhoods is considered as an obstacle to econometric identification of neighborhood effects. Since neighborhood effects emanate from characteristics of neighborhoods that have been themselves objects of choice, their characteristics are not independent of characteristics of those who have chosen to reside in them. In the canonical Manski (1993) framework, where individuals are subject to contextual effects, such as the characteristics of neighborhoods and their residents, and to endogenous social effects, such as decisions by their neighbors, dependence of between individuals' characteristics and those of their neighborhoods add a layer of complication to econometric identification. That is, individual decisions reflect at equilibrium the decisions of all other members of the community. A number of works on social interactions, including mainly theoretical works as Brock and Durlauf (2001), Durlauf (2004), Glaeser and Scheinkman (2001) and Moffitt (2001), have sought to obtain conditions under which endogenous effects may be separated from contextual social effects. These ideas also apply when the concept of the neighborhood is broadly construed to also include social groups, local communities, and political jurisdictions like cities and metropolitan areas. Individuals' equilibrium sorting into communities has of course been recognized by this literature. In fact, such early contributions as Evans, Oates and Schwab (1992) recognize the difficulty of econometric identification of local community-level effects and propose instrumental variable techniques to address it. More recently, Epple and Platt (1998) and Epple and Sieg (1999) address directly properties of equilibrium sorting into communities. Ioannides and Zabel (2008) and Ioannides and Zanella (2008) seek to identify endogenous neighborhood effects via individuals' choice of residential neighborhoods.

Nesheim (2002) embeds the problem in the explicit hedonic setting of Ekeland, Heckman and Nesheim (2004), who seek to clarify general conditions under which one may recover preferences and production relationships from "hedonic market" data. These are markets where goods with multiple attributes are valued. Examples abound and include workers with different skills, automobiles with different features, and houses that differ in terms of structural characteristics and attributes of the neighborhoods in which they lie. The Nesheim and Ekeland, Heckman and Nesheim solutions in their greatest generality are not amenable to explicit analytical solutions. In fact, the problem typically involves integration of a general partial differential equation that is rarely possible to do in closed-form.

In the remainder of this paper we fix ideas by presenting first Nesheim's formulation of a general endogenous sorting with contextual effects and income effects and his solution of its special case when income effects are excluded. Then we present our solution of his general model, with contextual effects as well as own income effects actually present.

2 Nesheim's Model of Sorting

In Nesheim (2002) individuals choose their location from among a continuum of locations that differ in terms of an attribute that individuals value, average neighborhood schooling among the adults. Individuals consume a unit of housing each and have preferences over non-housing and their own child's expected schooling. That is, an individual (household) allocates her income over non-housing consumption, $I - R(\ell)$, where $I, R(\ell)$, are, respectively, income and unit housing rent at location ℓ, $\ell \in R_+$.

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

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

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

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

Schooling production in location ℓ is defined via 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:

$$s_1 = (S(\ell))^{\eta_1} e^{\eta_2 z_1 + z_3 + z_5}, \tag{2}$$

where η_1, 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 independent of location. Average schooling at each location, $S(\ell)$, and the housing rent, $R(\ell)$, are both endogenous. They must be determined consistently with equilibrium sorting of households across locations.

The sorting equilibrium is defined in terms of: one, a mapping $F(z)$ that assigns household types to locations, $z \in Z$, $F(z) : Z \to R_+$; and two, a housing price function $R(\ell)$ that equilibrates the housing market in all locations $\ell \in R_+$. In other words, $S(\ell)$ is endogenous, being defined as the average schooling of parents who have chosen to locate in neighborhood ℓ, $S(\ell)$, $S(\ell) = \mathcal{E}\left[e^{z_1}|z \in F^{-1}(\ell)\right]$.

Individuals take the sorting equilibrium as given. The analytics are simplified because conditional on an individual's characteristics, location in the utility maximization problem may be represented via average schooling of parents in each neighborhood, $S(\ell)$. Therefore, under the assumption of monotonicity (which needs to be satisfied at equilibrium), the choice of location is equivalently handled via the choice of S. Consequently, the price function is equivalently defined as a function of the contextual effect, S, $p(S)$, that is instead of price of location, $R(\ell)$.

The special case of $\gamma = 0$ renders the first term in the utility function (1) linear in non-housing consumption, with the utility maximization problem becoming:

$$\max_S \left\{ e^{z_2} - p(S) + e^{z_4} S^{\eta_1} e^{\eta_2 z_1 + z_3} \mathcal{E}(e^{z_5}) \right\}. \tag{3}$$

The utility maximization problem is simplified enormously, but the impact of income on neighborhood choice is excluded. The first-order condition involves the equilibrium price function in terms of S :

$$-p_S(S) + \eta_1 \mathcal{E}(e^{z_5}) S^{\eta_1 - 1} e^{\eta_2 z_1 + z_3 + z_4} = 0. \tag{4}$$

Viewed alternatively, this condition designates a set of characteristics of individuals who choose expected neighborhood schooling S, given price $p(S)$. Clearly, it implies tradeoffs among such individuals' own schooling, their child's ability and the relative weight they assign to schooling. Rearranging (4) and taking logs yields:

$$\eta_2 z_1 + z_3 + z_4 = \ell n \left[p_S(S) \right] - \ell n [A_0 \eta_1] + (1 - \eta_1) \ell n S, \tag{5}$$

where $A_0 = \mathcal{E}(e^{z_5})$. If $z_5 \sim N(0, \sigma_5^2)$ and uncorrelated with all other z's, then $A_0 = e^{\frac{1}{2}\sigma_5^2}$. The RHS above is the marginal effect of average neighborhood schooling, neighborhood quality, measured in terms of S, and price on utility; the LHS is the marginal willingness to pay for neighborhood quality.

Average neighborhood schooling, that is, average schooling of parents who choose location S, may now be defined in terms of condition (5):

$$S = \mathcal{E} \left\{ e^{z_1} | \eta_2 z_1 + z_3 + z_4 = \ell n p_S(S) - \ell n (A_0 \eta_1) + (1 - \eta_1) \ell n S \right\}. \tag{6}$$

Once an unconditional multivariate normal distribution is assumed for the vector z, then it is elementary to derive an expression for the right hand side of (6) [Nesheim (2002), p. 21], the average schooling of individuals who choose location S, conditional on (5). The can be rewritten as a differential equation in $p(S)$.

2.1 Endogenous Contextual Effects and Hedonic Prices

Let $\Theta(S)$ denotes the one-dimensional index of neighborhood quality, the RHS of (5):

$$\Theta(S) \equiv \ell n[p_S(S)] - \ell n[A_0\eta_1] + (1-\eta_1)\ell n S,$$

and let $\phi^{\mathrm{T}} = (\eta_2, 0, 1, 1, 0)$. Then $\phi^{\mathrm{T}} z$, the LHS of (5), denotes an one-dimensional index of consumer willingness to pay for educational quality. If $z \sim N(\mu, \Sigma)$, then the log education of parents who choose location S is normally distributed:

$$\left(z_1 | \phi^{\mathrm{T}} z = \Theta(S)\right) \sim N\left(\widetilde{\mu}_1, \widetilde{\sigma}_1\right),$$

with mean and variance, according to standard theory, given by

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

where: $\widetilde{\rho}^2_1 = \frac{\left(\phi^{\mathrm{T}}\Sigma e_1\right)^2}{\sigma_{11}(\phi^{\mathrm{T}}\Sigma\phi)}$, $e^{\mathrm{T}}_1 = (1,0,0,0,0)$, 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 the expression for the expectation of e^{z_1}, conditional on the willingness to pay for neighborhood quality being equal to location quality, we have $S = \mathcal{E}\left\{e^{z_1} | \phi^{\mathrm{T}} z = \Theta(S)\right\} = e^{\widetilde{\mu}_1 + \frac{1}{2}\widetilde{\sigma}^2_{11}}$. Rewriting it yields:

$$S = L_0 p_S^{L_1} S^{L_1(1-\eta_1)}, \tag{7}$$

where the auxiliary variables L_0, L_1 are functions of behavioral parameters and of the parameters of the joint distribution of vector z, defined as follows:

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

Note that L_1 is the covariance between willingness to pay for neighborhood quality and log education of parents.

Integrating differential equation Equ. (7) and using the assumption that locations with average education equal to 0 are priced at 0, $p(0) = 0$, yields

$$p(S) = L_0^{-\frac{1}{L_1}} \left(\eta_1 + \frac{1}{L_1} \right)^{-1} S^{\eta_1 + \frac{1}{L_1}}. \tag{8}$$

This is the equilibrium price function, the hedonic price of housing, as a function of average neighborhood education of parents, the contextual effect.

The interpretation of this solution is interesting. The elasticity of the hedonic price with respect to neighborhood quality is equal to the elasticity of schooling with respect to quality, η_1, plus the inverse of the covariance between willingness to pay for neighborhood quality and log education of parents, $\frac{1}{L_1}$. The larger is η_1, the larger percentage price differentials must be to segregate individuals into their preferred locations. The larger the covariance between willingness to pay for neighborhood quality and parental education, 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 differentials are required to be to maintain segregation of households at equilibrium. The equilibrium price function (8) is convex (concave), if $\eta_1 + \frac{1}{L_1} > (<) 1$. However, for the solution to be meaningful, $\eta_1 + \frac{1}{L_1} > 0$.

This theory yields three specific testable implications. First is the educational production function along the lines of (2), a relationship between schooling outcomes, observable inputs in neighborhoods, including contextual effects and unobservable individual and family characteristics. This is typically postulated in the social interactions literature and is testable, in principle. Second is the first-order condition for the optimal location, (4), which is the hedonic demand equation, an individual's marginal valuation of location, defined in terms of average schooling of parents, as a function of individual characteristics. And third, the actual equilibrium housing price function, (8). These three relationships depend on the same set of deep parameters, namely the parameters of the utility and educational production functions and of the joint distribution function of parental schooling, of parental income, of child's ability in school, log of a preference parameter that enters a child's schooling outcome, relative valuation of child's schooling, and a stochastic shock to schooling outcome. Furthermore, the distributions of education and of income of parents who choose neighborhood education S are quite straightforward to derive.

The index of neighborhood quality depends on S only: $\Theta(S) \equiv -\frac{1}{L_1}\ell n L_0 - \ell n(A_0\eta_1) + \frac{1}{L_1}\ell n S$. So, from (5) we have:

$$S = A_0^* e^{L_1(\eta_2 z_1 + z_3 + z_4)},\ A_0^* \equiv L_0^{-\frac{1}{L_1}}(A_0\eta_1)^{L_1}.$$

Note that the log of average neighborhood schooling chosen is linear in $L_1(\eta_2 z_1 + z_3 + z_4)$, where the effect of individual characteristics is moderated by L_1. However, the larger the covariance L_1 between willingness to pay for neighborhood quality and parental education, the greater the elasticity of neighborhood education with respect to parental schooling. That is, the larger is L_1, 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. We see below that this result is modified substantially in the presence of income effects.

Also interesting is the relationship between the log of a child's education and household characteristics. That is:

$$\ell n s_1 = a_0 + \eta_2(1 + \eta_1 L_1)z_1 + (1 + \eta_1 L_1)z_3 + \eta_1 L_1 z_4 + z_5, \tag{9}$$

where a_0 is a function of parameters. If covariance L_1 and/or η_1 is very large, it is possible that the log of children's education increase more than proportionately with that of their parents.

The hedonic price literature typically estimates an *arbitrary* functional form for the relationship between the (marginal) valuation for neighborhood amenities as a function of observables and unobservables. However, it is a key implication of Nesheim's theory that individual utility maximization exactly determines the equilibrium price, as a function of contextual characteristics,and this in turn determines sorting. A typical estimation approach would be to estimate a system of equations for individual education outcomes and average parental education by neighborhood. As Nesheim (2002) shows in detail, this model is not entirely identified, although groups of parameters may be identified by means of data on observable educational outcomes as function of parental education, neighborhood school quality, and income.

3 The Case of Income Effects

Nesheim (2002) claims that when $\gamma \neq 0$, the counterpart of the differential equation (7), that the hedonic price function must satisfy, does not lend itself to an analytical solution and may only be solved numerically.[1]

We modify the underlying assumptions of Nesheim (2002) by assuming that component z_2 of the vector of individual characteristics z is the *level* of parental income, *not* its log. This modification does break the symmetry of Nesheim's setting but pays off nicely in terms of an analytical solution for $p(S)$. The first order condition for maximizing utility (1) is now, instead of (4):

$$-e^{-\gamma[z_2-p(S)]}p_S(S) + A_0\eta_1 S^{\eta_1-1}e^{\eta_2 z_1+z_3+z_4} = 0. \tag{10}$$

Rearranging and taking logs yields the counterpart of (5):

$$\eta_2 z_1 + \gamma z_2 + z_3 + z_4 = \Theta(S), \tag{11}$$

where:

$$\Theta(S) \equiv \gamma p(S) + \ell n\left[p_S(S)\right] - \ell n(A_0\eta_1) + (1-\eta_1)\ell n S. \tag{12}$$

This more general formulation differs because the willingness to pay includes an effect of parental income, γz_2, and the neighborhood quality index $\Theta(S)$ includes the term $\gamma p(S)$. Defining the vector of parameters φ now as

$$\varphi^{\mathrm{T}} = (\eta_2, \gamma, 1, 1, 0) \tag{13}$$

and using it in place of ϕ in the derivations and formulas, we now have instead of (7) the differential equation:

$$S = L_0\left(p_S(S)\right)^{L_1} S^{L_1(1-\eta_1)}e^{\gamma p(S)}. \tag{14}$$

[1] Interestingly, Nesheim also shows that this more general model is econometrically fully identified.

Integrating this equation and using the initial condition $p(0) = 0$, yields the equilibrium price function:

$$p(S) = \frac{L_1}{\gamma} \ell n \left[1 + \frac{\gamma}{L_1} L_0^{-\frac{1}{L_1}} \left(\eta_1 + \frac{1}{L_1} \right)^{-1} S^{\eta_1 + \frac{1}{L_1}} \right]. \tag{15}$$

We see the properties of this more general result for the price function by differentiating (15). The equilibrium price function is increasing in S. Its curvature is determined by the sign of its second derivative:

$$p_{SS}(S) = L_0^{-\frac{1}{L_1}} S^{\eta_1 + \frac{1}{L_1} - 2} \frac{\eta_1 + \frac{1}{L_1} - 1 - \frac{\gamma}{L_1} L_0^{-\frac{1}{L_1}} \left(\eta_1 + \frac{1}{L_1} \right)^{-1} S^{\eta_1 + \frac{1}{L_1}}}{\left[1 + \frac{\gamma}{L_1} L_0^{-\frac{1}{L_1}} \left(\eta_1 + \frac{1}{L_1} \right)^{-1} S^{\eta_1 + \frac{1}{L_1}} \right]^2}. \tag{16}$$

Therefore, the equilibrium price function is convex for low values of S and up to the threshold point, concave thereafter, and provided that $\eta_1 + \frac{1}{L_1} > 1, \ \gamma > 0$. Here we see that a key property of social interactions models is reaffirmed, if this condition is satisfied: the equilibrium price is a sigmoid function of average neighborhood quality. In view of the definition of L_1 above, the covariance between willingness to pay for neighborhood quality and parental education, this condition involves restrictions in terms of parameters and of the stochastic structure:

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

As expected, the parameters describing the distribution of income also enter the equilibrium price function. Both solutions (7) and (14) underscore an important property of hedonic price theory. While an individual's marginal valuation of neighborhood quality, the hedonic demand equation (11), now does depend on the individual's income, the equilibrium price function (14) depend, via the functions L_0, L_1, only on the statistics of the income distribution and its joint distribution with the other characteristics of interest. Furthermore, the theory allows us to relate the statistics of income distribution, conditional on average neighborhood schooling, to its equilibrium price.

How does our solution with income effects compare with that of Neisheim's? First, we note that for small S, (15) implies (8). Second, for large S, the second term within the brackets in the RHS of (15) dominates the first, thus implying:

$$p(S) \approx \frac{L_1}{\gamma} \ell n \left[\frac{\gamma}{L_1} L_0^{-\frac{1}{L_1}} \left(\eta_1 + \frac{1}{L_1} \right)^{-1} \right] + \frac{1 + \eta_1 L_1}{\gamma} \ell n S. \tag{17}$$

In this case, the elasticity of equilibrium price with respect to neighborhood quality is increasing in $\eta_1 L_1$ and decreasing in the coefficient of absolute risk aversion γ with respect to non-housing consumption. In contrast, this elasticity in the absence of income effects, $\eta_1 + \frac{1}{L_1}$, is quite different and reflects an additional parameter, γ.

The optimal value of average neighborhood schooling for an individual with characteristics (z_1, z_2, z_3, z_4) may be determined from (10), by using the solution (15). This yields:

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

where $\tilde{A}_0 \equiv A_0 \eta_1 L_0^{\frac{1}{L_1}}$. The choice of neighborhood for an individual with characteristics (z_1, z_2, z_3, z_4) is implicitly determined from the above equation. Given a solution for S, $S = \mathcal{S}(z_1, z_2, z_3, z_4)$, then the child's education follows from (2), $\ell s_1 = \eta_2 z_1 + z_3 + z_5 + \eta_1 \mathcal{S}$. The dependence of a child's education on parental income is brought in via $\mathcal{S}$.

If $L_1 > 1$, the covariance between the willingness to pay for neighborhood education quality and the income of parents, is greater than 1, then the LHS of the above equation is increasing monotonically in S, and therefore there exists a unique solution for any given value of the willingness to pay neighborhood for education quality $\eta_2 z_1 + \gamma z_2 + z_3 + z_4$, with respect to which it would be increasing. In particular, individuals with larger incomes choose higher average neighborhood quality. And holding neighborhood quality constant, incomes are negatively correlated with individuals' own education and with their children's schooling ability [*c.f.* Epple and Platt (1998)]. Individuals with more own schooling choose neighborhoods with higher average schooling, which in turn implies higher schooling for their children.

The magnitude of L_1 does affect the sensitivity of S with respect to the willingness to pay neighborhood for education 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. They have with different properties. For example, the higher value 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 way to see the consequences of this approach is to consider the distribution of income, conditional on location choice S. The income of people who choose S is normal:

$$\left(z_2 | \varphi^{\intercal} z = \Theta(S)\right) \sim N\left(\tilde{\mu}_2, \tilde{\sigma}_{22}\right),$$

with mean and variance

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

where $\tilde{\rho}_2^2 = \frac{\left(\varphi^{\intercal}\Sigma e_2\right)^2}{\sigma_{22}(\varphi^{\intercal}\Sigma\varphi)}$, denotes the square of the correlation coefficient between log education of parents and willingness to pay for neighborhood quality, and the solution (15) is used in the expression for $\Theta(S)$, in (12). In this case, function $\Theta(S)$ becomes:

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

A sufficient condition for the mean income of parents who choose S to increase with S is $L_1 - 1 > 0$. Similarly, the conditional distribution of education of parents who choose has a tractable dependence on their own education.

4 Relationship to Previous Literature

The paper aims at contributing to a recent literature on the influence on schooling outcomes for children from the educational attainment of their adult neighbors. 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 non-linearities in the general relationship, whose special case is estimated by Kremer (1997), may alter his key results.

The present paper is most closely related to Nesheim (2002), whose pioneering contribution emphasizes the role of housing price in residential sorting and derives theoretically the properties of associated hedonic housing functions. Unfortunately, Nesheim's most general model may not be solved analytically, and the model that is actually solved does not allow for income effects in neighborhood choice. This affects critically the functional form of the hedonic price.

Graham (2008)[2] emphasizes the econometric identification of neighborhood exernalities in a model where schooling is a function of parental and neighborhood schooling, and where adult income is a function of own schooling. Human capital acquisition ability and labor market ability may co-vary. Graham assumes that the total ability effect is jointly distributed with parental human capital and 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. In his behavioral model, parents care for own consumption and for expected consumption by their children, may borrow, and purchase housing. Housing "buys" access to neighborhood amenities, and its hedonic price is thus a function of neighborhood quality. Intertemporal optimization yields an Euler-like equation, which links the hedonic price of housing to parental schooling, neighborhood schooling and total ability. Graham's equation is the counterpart of Equ. (12) above. Its integration yields a hedonic price as an isoelastic function of neighborhood schooling. This is the exact counterpart of Nesheim's analytic solution, reproduced here as Equ. (8), and is thus clearly less general than our derivation of the hedonic price function (15) from a model that also allows for income effects on individuals' location decisions.

[2]Graham and I were unaware of each other's work. When he generously shared with me his Graham (2008) working paper, I reciprocated by sharing with him the March 3, 2008 version of Ioannides (2008).

5 Conclusions

This paper extends results by Nesheim (2002) in an important direction by allowing for income effects in neighborhood choice. Since individuals value neighborhood schooling as an input to schooling attained by their children, they self-select into neighborhoods. That is, parents' choice of neighborhood and therefore of their children's education is influenced by their own income. Therefore, assessing the impact of neighborhood schooling on children's schooling requires accounting for sorting bias.

For an individual's child, schooling outcome is assumed to be a function of average parental schooling in the neighborhood, own schooling, own income, the child's ability, the weight of schooling in her preferences and a random factor. Individuals' utility maximizing choices imply a hedonic price of housing which mediates sorting of individuals into neighborhoods at equilibrium. The resulting average parental schooling in neighborhoods, the single quality attribute of each neighborhood in the model, is consistent with individuals' decisions.

We may use the model to study the dynamics of the co-evolution of individual and neighborhood schooling. The fact that neighborhood choice is affected by income has rich implications. The intertemporal evolution of schooling may be non-linear. Linking income with adult schooling would allow us to study the effect on the intertemporal evolution of the income distribution of neighborhood schooling as an input to the production of schooling.

6 References

Brock, William A. and Steven N. Durlauf. 2001. "Interactions-based Models." *Handbook of Econometrics* 5: 3297 - 3380, James Heckman and Edward Leamer, Eds. Amsterdam: North-Holland.

Brooks-Gunn, Jeanne, Greg J. Duncan, Pamela Kato Klebanov and Naomi Sealand. (1993). "Do Neighborhoods Influence Child and Adolescent Development?" *American Journal of Sociology* 99 (2):353 – 395.

Durlauf, Steven N. 2004. "Neighborhood Effects." *Handbook of Regional and Urban Economics* 4: 2173 - 2242, J.Vernon Henderson and Jacques-Francois Thisse, Eds. Amsterdam: North- Holland.

Ekeland, Ivar, James J. Heckman, and Lars Nesheim. 2004. "Identification and Estimation of Hedonic Models." *Journal of Political Economy* 112 (S1): S60 - S109.

Epple, Dennis and Glen J. Platt. 1998. "Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Incomes." *Journal of Urban Economics* 43 (1): 23 – 51.

Epple, Dennis and Holger Sieg. 1999. "Estimating Equilibrium Models of Local Jurisdictions." *Journal of Political Economy* 107 (4): 645 – 681.

Evans, William N., Wallace E. Oates, and Robert M. Schwab. 1992. "Measuring Peer Group Effects: a Study of Teenage Behavior." *Journal of Political Economy* 100(5): 966 – 991.

Glaeser, Edward L., and José A. Scheinkman. 2001. "Measuring Social Interactions." in *Social Dynamics*: 83 – 130, Steven N. Durlauf and Peyton Young, Eds. Cambridge, MA: MIT Press.

Graham, Bryan S. 2008. "On the Identification of Neighborhood Externalities in the Presence of Endogenous Neighborhood Selection." Department of Economics, University of California, Berkeley, June 26.

Ioannides, Yannis M. 2003. "Empirical Nonlinearities and Neighborhood Effects in the Intergenerational Transmission of Human Capital." *Applied Economics Letters,* 10, 535–539.

Ioannides, Yannis M. 2008. 'Neighborhood Effects and Housing." written for *Handbook of Social Economics,* Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, eds., Elsevier, forthcoming, version March 3.

Ioannides, Yannis M. and Jeffrey E. Zabel. 2008. "Interactions, Neighborhood Selection and Housing Demand." *Journal of Urban Economics* 63(1): 229 – 252.

Ioannides, Yannis M. and Giulio Zanella. 2008. "Searching for the Best Neighborhood: Mobility and Social Interactions." working paper, Department of Economics, Tufts University, April.

Kremer, Michael. 1997. "How Much Does Sorting Increase Inequality?" *Quarterly Journal of Economics* 112(1): 115 – 139.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: the Reflection Problem." *Review of Economic Studies* 60 (3): 531 - 542.

Moffitt, Robert A. 2001. "Policy Interventions, Low-level Equilibria and Social Interactions." in *Social Dynamics*: 45 - 82 Steven N. Durlauf and H. Peyton Young, Eds. Cambridge, MA: MIT Press.

Nesheim, Lars. 2002. "Equilibrium Sorting of Heterogeneous Consumers across Locations: Theory and Empirical Implications." CEMMAP Working Paper CWP08/02, University College London.