Household Sorting and Neighborhood Formation*

Dennis Epple

Carnegie Mellon University and NBER

Michael Peress

University of Rochester

Holger Sieg

Carnegie Mellon University and NBER

We develop a new model of household sorting in a system of residential neighborhoods. We show that this model is partially identified without imposing parametric restrictions on the distribution of unobserved tastes for neighborhood quality and the shape of the indirect utility function. The proof of identification is constructive and can be used to derive a new semiparameteric estimator. Our empirical application focuses on residential choices and housing demand in a system of neighborhoods in the Pittsburgh metropolitan area. We find that there are significant differences in the observed sorting of households with and without children. In particular, households with children exhibit more stratification by income than households without children.

*This research was motivated by discussions with James Heckman during the the 2002 CAM workshop on "Characteristics Models" at the University of Copenhagen. We also thank Pat Bajari, Pat Bayer, Lanier Benkhard, Steve Berry, Herman Bierens, Martin Browning, Jan Bruckner, Andrew Chesher, Phil Haile, Lars Nesheim, Charles Manski, Rosa Matzkin, Antonio Merlo, Ariel Pakes, Richard Romano, John Rust, Matt Shum, Chris Taber, Quang Voung, Ed Vytlacil, and participants of seminars at a CEPR conference on Urban Economics at LSE, the SITE Workshop on Structural Estimation, the ASSA Winter Meetings in San Diego and Boston and workshops at Arizona, Berkeley, Chicago, Florida, Johns Hopkins, Minnesota, Northwestern, Penn State, and Yale for comments and suggestions. Financial support for this research is provided by the NSF under Grant No. 0111630, the MacArthur Foundation, and the Alfred P. Sloan Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, the MacArthur Foundation, or the Alfred P. Sloan Foundation.

1 Introduction

Research over the past several years has led to development of models characterizing equilibrium in a system of local jurisdictions. An important insight from these models is that plausible single-crossing assumptions about preferences generate strong predictions about the equilibrium distribution of households across communities and neighborhoods. More recently, research has focused on devising empirical strategies, which can be used to estimate the parameters of these models, evaluate their goodness of fit, and to derive strategies for applied welfare analysis based on these models.[1] We develop a new sorting model that uses mixtures of distributions to characterize observed and unobserved heterogeneity among households. This approach allows us to model differences in discrete as well as continuous characteristics of households. The resulting equilibrium model can, therefore, be viewed as a mixture of hierarchical models of the type considered in Epple and Sieg (1999).

We then adopt a nonparametric approach to study identification our model. We allow households to differ in their incomes, their tastes for neighborhood quality, and various other characteristics such as family structure. Since the distribution of tastes for neighborhood quality is inherently unobservable, we do not impose functional form assumptions on the distribution of tastes.[2] One objective of the analysis is then to derive general conditions that allow us to nonparametrically identify the distribution of household characteristics and the indirect utility function of households based on the observed equilibrium outcomes.

We first consider the case in which the utility function is known. We show that the discreteness of the choice set imposes limits to identification.[3] We find that it is possible to

[1]Parametric versions of the models considered in this paper have provided new insights into household behavior as discussed, for example, in Epple and Sieg (1999),Epple, Romer, and Sieg (2001), Wu and S. (2003), Sieg, Smith, Banzhaf, and Walsh (2004), Calabrese, Epple, Romer, and Sieg (2006), Ferreyra (2007), and Walsh (2007). See also Nesheim (2001), Bajari and Kahn (2004), Bayer, McMillan, and Reuben (2004) Bayer, Ferreira, and McMillan (2007), and Ferreira (2005) for related empirical approaches which are based on more traditional discrete choice models or hedonic frameworks.

[2]Our research is thus similar to recent work by Ekeland, Heckman, and Nesheim (2004), Heckman, Matzkin, and Nesheim (2004) and Bajari and Benkard (2005) on identification and estimation of hedonic models. It is also closely related to Blundell, Browning, and Crawford. (2003) who discuss nonparametric tests of revealed preference models. There are also some similarities to the literature on selection in the Roy model as discussed, for example, in Heckman and Honore (1990).

[3]Point identification cannot be achieved in many econometric applications. In that case, attention nat-

nonparametrically identify a finite number of points of the distribution of tastes conditional on income for each household type. These points correspond to the points on the boundary between adjacent neighborhoods. For points that are not on the boundary loci we can only provide lower and upper bounds for the distribution. These bounds become tighter as the number of differentiated neighborhoods in the application increases. Joint non- or semi-parametric identification of the distribution of household characteristics and the indirect utility function is more difficult to establish. We show that we can provide bounds for the indirect utility function and the distribution of tastes. We thus conclude that the model is partially nonparametrically identified.[4]

Our proofs of identification are constructive and give rise to algorithms for estimating the functions of interest or placing bounds on important parameters. Our sorting model implies that the (unobserved) quality of the neighborhoods should be monotonically increasing in price or the price rank of the neighborhood. Moreover, this function must also have a sufficient degree of curvature to guarantee that the differences in qualities are large enough given the observed differences in prices. We can nonparametrically estimate a function which links the observed measure of quality of the neighborhood to the observed price rank of the neighborhood using recent innovations in nonparametric estimation which impose monotonicity and curvature constraints on the underlying function.[5] We then derive a testing procedure which allows us to determine which set of parameter values of the indirect utility functions are admissible, i.e. we test for which parameters the shape restrictions are valid. For each admissible indirect utility function, we can estimate the joint distribution of income and tastes.[6]

The empirical analysis focuses on residential and housing choices in the Pittsburgh metropolitan area. Our framework allows neighborhoods to differ in quality (local public

urally shifts to characterizing informative bounds on the parameters of interest. Some recent examples are Manski (1997) and Tamer (2003).

[4]For a general discussion of partial identification see Imbens and Manski (2004) and Chernozhukov, Imbens, and Newey (2006).

[5]See Matzkin (1994) and Lewbel and Linton (2003) for a discussion of nonparametric estimators that impose shape restrictions.

[6]Our estimation approach thus differs significantly from share inversion estimators discussed in Berry (1994) and Berry, Levinsohn, and Pakes (1995).

goods such as public education and protection from crime) and households may choose to purchase different quantities of housing in each neighborhood. We find that there are significant differences in the observed sorting of households with and without children. In particular, households with children exhibit more stratification by income than households without children. Low-income households with children have lower tastes for local public goods and amenities than similar households without children. The opposite is true for high-income households. Households with children choose to reside in larger homes in cheaper neighborhoods.

The rest of the paper is organized as follows. Section 2 discusses the household demand model. Section 3 discusses identification. Section 4 develops a new semiparametric estimator motivated by our identification results. Section 5 discusses the data used in this paper. Section 6 presents the main empirical findings. Section 7 offers some conclusions and discusses future research.

2 Household Sorting and Residential Choices

We consider a metropolitan area with a finite number of household types that differ in their endowed income, y, and in a taste parameter, α, which reflects the household's strength of preferences for neighborhood quality. Each household type i captures some observed discrete differences in the population such as family structure and occurs with probability P_i. The continuum of households conditional on type i is implicitly described by the joint distribution of α and y, denoted by $F_i(\alpha, y)$.

Assumption 1 *The joint distribution of income and tastes $F_i(\alpha, y)$ is continuous with support $S \subseteq R^2_+$ and joint density $f_i(\alpha, y)$, for i=1,..,I.*

A household of type i with taste parameter α and income y is referred to as a triple (α, y, i).

A household has preferences defined over the quality of the neighborhood, g, the quantity of housing consumed, q, and a composite private good, b.[7]

Assumption 2 *The preferences of a household are represented by a utility function, $U(\alpha, g, q, b)$ that is twice differentiable in its arguments and strictly quasi-concave in g, q, and b.*

Conditional on choosing a neighborhood, thereby choosing g, the household determines the optimal amount of housing q by choosing

$$\begin{aligned} &\max_{(q,b)} \; U(\alpha, g, q, b) \\ &s.t. \;\; p\, q \;=\; y \;-\; b \end{aligned} \tag{1}$$

where p denotes the gross-of tax price of housing. There are J different neighborhoods in the metropolitan area. Without loss of generality we assume that $g_1 < ... < g_J$ which of course implies that $p_1 < .... < p_J$.

It is convenient to represent the preferences of a household that lives in neighborhood j using the indirect utility function, $V(\alpha, y, g_j, p_j)$.

$$V(\alpha, g_j, p_j, y) \;=\; U(\alpha, g_j, q(p_j, y, \alpha), y - p_j q(p_j, y, \alpha))$$

Households choose the neighborhood which maximizes (indirect) utility:

$$\max_{d_1,\ldots,d_J} \; \sum_{j=1}^{J} d_j \; V(\alpha, g_j, p_j, y)$$

subject to the constraint that $\sum_{j=1}^{J} d_j = 1$ and $d_j \in \{0, 1\}$. Note that the "error term" of the model (α) enters into the indirect utility function in a non-additively separable way. As such the model does not give rise to a standard random utility model with additive error

[7]We are thus assuming that households have the same utility function conditional on tastes, i.e. $U_i = U$ for all i. It is straightforward to extend the analysis in this paper to allow for differences in $U_i(\cdot)$.

terms.[8]

To characterize the sorting of households in this model, it is useful to impose additional assumptions on the indirect utility function. Consider the slope of an "indirect indifference curve" in the (g_j, p_j)-plane:

$$M(\alpha, y, g_j, p_j) \;=\; \frac{dp_j}{dg_j}\Big|_{V=\bar{V}} \tag{2}$$

We assume that the indirect utility function satisfies standard single-crossing conditions in the (g_j, p_j)-plane.

Assumption 3 *For given α, $M(\cdot)$ is monotonically increasing in y. For given y, $M(\cdot)$ is strictly monotonically increasing in α.*

Note that non-additive separability of α is a necessary (but not sufficient) condition for assumption 3 to hold.

Let C_j denote the set of households that choose neighborhood j:[9]

$$C_j \;=\; \{(\alpha, y) \,|V(\alpha, y, g_j, p_j) \geq \max_{i \neq j} V(\alpha, y, g_i, p_i)\} \tag{3}$$

The share of households of type i that live in neighborhood j is given by:

$$n_{ij} \;=\; \int_{C_j} f_i(\alpha, y) \;\, d\alpha \; dy \tag{4}$$

Summing over all discrete types yields the total size of neighborhood j:

$$n_j \;=\; \sum_{i=1}^{I} n_{ij} \; P_i \tag{5}$$

[8]The structure of the discrete part of the choice model is similar to a non-parametric ordered probit model. Han (1987) and Kahn (2001) allow the error term to enter as a separate argument into an unknown parametric function.

[9]The set of households that are indifferent have measure zero.

Consider an allocation in which each neighborhood has a positive size. For such an allocation to be an equilibrium – there must be an ordering of neighborhoods, $\{(g_1, p_1), ..., (g_J, p_J)\}$, with $g_1 < .... < g_J$ such that:

1. **Boundary Indifference:** There exists a set of households that are indifferent between two "adjacent" neighborhoods. This set is characterized by the following expression:

$$R_j \;=\; \{(\alpha, y) \mid \; V(\alpha, g_j, p_j, y) \;=\; V(\alpha, g_{j+1}, p_{j+1}, y)\} \qquad j = 1, ..., J-1 \tag{6}$$

2. **Stratification:** Let $\alpha_j(y)$ be the implicit function defined by equation (6).[10] Then, for each level of income y, the households that live in j consist of those with tastes, α, given by:

$$\alpha_{j-1}(y) \;\leq\; \alpha \;\leq\; \alpha_j(y) \tag{7}$$

3. **Ascending Bundles:** Consider two neighborhoods i and j such that $p_i \;>\; p_j$. Then $g_i > g_j$ if and only if $\alpha_i(y) \;>\; \alpha_j(y)$. In particular, boundary indifference loci do not intersect.

As we will see below, these necessary conditions of equilibrium are important in establishing identification of the model.[11]

Most of the previous empirical literature in urban economics also assumes that the indirect utility function satisfies the following separability assumption:[12]

Assumption 4 *The indirect utility function is additively separable and hence can be written*

[10]Define $\alpha_0(y) = 0$ and $\alpha_J(y) = \infty$.

[11]A formal proof of the results above are given in Epple and Sieg (1999).

[12]This assumption is also invoked by almost all computational general equilibrium studies such as Epple and Romer (1991), Nechyba (1997), Fernandez and Rogerson (1998), Nechyba (2000, 2003), Schmidheiny (2006), and Rothstein (2006).

as:

$$V(\alpha, y, g, p) \;=\; \alpha\, V^g(g) \;+\; V^b(y, p) \tag{8}$$

Using Roy's Identity, the demand functions are given by:

$$q(p, y) \quad = \quad -\frac{\partial V^b / \partial p}{\partial V^b / \partial y} \tag{9}$$

A direct consequence of assumption 4 is that the demand function above does not depend on α and g.[13]

Separability does not imply that the unconditional demand for housing does not depend on neighborhood characteristics. It just means that, conditional on neighborhood choice, neighborhood characteristics only affect the demand for housing if these neighborhood amenities are capitalized in the housing prices. This fact is largely supported by the empirical literature. Separability rules out a household choosing to buy a larger house just because it happens to live in a high-crime neighborhood. Almost all previous empirical and computational studies of local public good provision and residential sorting in in local housing markets are either implicitly or explicitly based on this type of separability assumption.

3 Identification

The nature of the identification problem is to determine whether it is possible to identify the indirect utility function $V_0(\alpha, y, g, p)$ and the joint distributions of income and tastes $\{F_{i0}(\alpha, y)\}_{i=1}^{I}$ given the observed outcomes. Identification depends largely on the information set that is available to the econometrician. We assume that the econometrician observes the following outcomes:

[13] Separability of the indirect utility function is admittedly a strong assumption. Dubin and McFadden's (1984) consider a parametric model with non-separable preferences.

Assumption 5 *For every neighborhood the econometrician observes:*

- *the share of households of type i that live in neighborhood j n_{ij}, i=1,...,I,*
- *the joint density of income and housing of each household type i=1,...,I,*

as well as prices, p_j and neighborhood qualities g_j.

We thus consider identification and estimation of this class of models based on aggregate or at least grouped data. These types of data are typically available from the U.S. Census and other publicly available sources.[14] We also follow Ekeland et al. (2004) and restrict attention to identification based on variation in one metropolitan housing market.[15]

First we consider the case in which $V(\cdot)$ is known to the econometrician.[16] Notice that knowledge of $V(\cdot)$ implies that the econometrician knows the boundary indifference loci $\alpha_j(y)$. The first result states that we can identify $J-1$ points of the conditional distribution of $F_i(\alpha|\, y)$ for each household type.

Proposition 1 *If the indirect utility function is known, one can identify $J-1$ points of $F_i(\alpha|\, y)$. These points correspond to the values of α implied by the $J-1$ boundary indifference loci, and are given by $\alpha = \alpha_j(y)$.*

Proof:

Note that the joint density of (α, y) of household type i that lives in neighborhood j is given by:

$$f_{ij}(\alpha, y) = \begin{cases} \frac{f_i(\alpha,y)}{n_{ij}} & \text{if} \quad (\alpha, y) \in C_j \\ 0 & \text{if} \quad (\alpha, y) \notin C_j \end{cases} \tag{10}$$

[14]For example, individual level data on residential choices at the local level are only available through Census Research Centers

[15]If we observe equilibria of the same market at successive points of time or if we observe multiple markets at one point of time, then additional sources for identification are possible. These approaches typically impose restrictions of how preferences are allowed to vary across markets. Some results for this case are available upon request from the authors.

[16]With a slight abuse of notation, we sometimes suppress the subscript 0 that denotes the model under which the data are generated. Similarly we do not use different symbols for marginal and joint distributions.

Hence the marginal density of income of households of type i that lives in j is given by:[17]

$$\begin{aligned} f_{ij}(y) &= \int_{C_j} f_{ij}(\alpha, y)\, d\alpha \\ &= \frac{f_i(y)}{n_{ij}} \int_{\alpha_{j-1}(y)}^{\alpha_j(y)} f_i(\alpha|y)\, d\alpha \\ &= \frac{f_i(y)}{n_{ij}}\, [F_i(\alpha_j(y)|\, y) - F_i(\alpha_{j-1}(y)|\, y)] \end{aligned} \tag{11}$$

Rearranging terms such that observables are on the right hand side of the equation yields for the first neighborhood:

$$F_i(\alpha_1(y)|\, y) \;=\; \frac{f_{i1}(y)}{f_i(y)}\, n_{i1} \tag{12}$$

and all other neighborhoods $j > 1$:

$$F_i(\alpha_j(y)|\, y) \;=\; \frac{\sum_{k=1}^{j} n_{ik} f_{ik}(y)}{f_i(y)} \tag{13}$$

The right-hand-sides of equations (12) and (13) are observed in the data by assumption 5. Hence, we can identify $J-1$ points of the conditional distribution function of α given y for each type i. These points correspond to the values of $\alpha_j(y)$, $j = 1, 2, ..., J-1$. Q.E.D.

The results in Proposition 1 are similar to results found in the econometric literature that considers nonparametric estimation of the ordered probit models and other semi-parametric discrete choice models.[18] The main difference here is that we consider identification based on the type of aggregate data that typically available to researchers in urban economics. But as in the semi-parametric discrete choice literature, we find that the discreteness of the choice set implies that we can arbitrarily transform the distribution of α given y on the intervals $(\alpha_{j-1}(y),\ \alpha_j(y))$ without affecting the sorting of households among neighborhoods in equilibrium, as long as the transformed distribution has the correct values at the boundaries. As a consequence the conditional distribution of tastes given income is not identified

[17] We assume that the support of y is the same for all neighborhoods.

[18] See Powell (1994) for a survey of the semi-parametric literature.

in the interior of these intervals.

While we do not obtain point identification of $F_i(\alpha|\ y)$ for points which are not on the boundary loci of the model, the monotonicity of the distribution function allows us to construct bounds for these function values. Let $\underline{F}_i(\alpha|\ y)$ ($\overline{F}_i(\alpha|\ y)$) denote the lower (upper) bound. For any value of α such that $\alpha_j(y) < \alpha < \alpha_{j+1}(y)$ we then obtain:

$$\underline{F}_i(\alpha|\ y) \ = \ F_i(\alpha_j(y)|\ y) \ \leq \ F_i(\alpha|\ y) \ \leq \ F_i(\alpha_{j+1}(y)|\ y) \ = \ \overline{F}_i(\alpha|\ y) \tag{14}$$

If there are many neighborhoods, each with a small size, we would expect that the difference between $\alpha_j(y)$ and $\alpha_{j+1}(y)$ will be small for most adjacent neighborhoods. We thus conclude that the bounds for the conditional distribution of tastes are likely to be informative in applications with large choice sets.

We have assumed that $V(\cdot)$ is known to the econometrician. We now consider the problem of jointly identifying $V(\cdot)$ and $\{F_i(\cdot)\}_{i=1}^{I}$. Assumption 5 implies that we observe the joint distribution of income and housing for each neighborhood. If the function $V^b(y,p)$ satisfies standard integrability conditions, identification of $V^b(y,p)$ is limited only by the fact that we observe a finite number of neighborhoods. Identification of $V^g(g)$ is more problematic. The main problem is that any sub-utility function $V^g(g)$ that yields an indirect utility function satisfying the single-crossing conditions and that implies boundary indifference loci that do not intersect and hence satisfy:

$$\alpha_{j+1}(y) \quad > \quad \alpha_j(y) \qquad \forall j \tag{15}$$

is consistent with the observed outcomes. As a consequence, we have the following result:

Proposition 2 *For any sub-utility functions $V^b(p,y)$ and $V^g(g)$ such that*
(i) $V(\alpha,y,g,p) = \alpha V^g(g) + V^b(p,g)$ satisfies assumptions 2, 3, and 5;
(ii) $V(\alpha,y,g,p)$ implies the same demand function as the true indirect utility function $V_0(\alpha,y,g,p)$;
(iii) boundary loci do not intersect, i.e. $\alpha_{j+1}(y) > \alpha_j(y)\ \forall j$,

there exist a set of distribution $\{F_i(\alpha, y)\}_{i=1}^I$ *such that the observed sorting of households is identical to the one obtained for the true model* $V_0(\alpha, y, g, p)$ *and* $\{F_{i0}(\alpha, y))\}_{i=1}^I$.

Proof:

Consider an indirect utility function $V_a(\alpha, y, g, p)$ that satisfies condition (i). Let $\alpha_j^0(y)$ and $\alpha_j^a(y)$ denote the boundary indifference loci that correspond to $V_0(\cdot)$ and $V_a(\cdot)$ respectively. Condition (iii) implies that

$$\alpha_{j+1}^a(y) \quad > \quad \alpha_j^a(y) \quad \forall j \tag{16}$$

Define the conditional distribution of α, $F_{ia}(\alpha|\, y)$ for the relevant points on the boundary loci as follows:

$$F_{ia}(\alpha_j^a(y)|\, y) \quad \equiv \quad F_{i0}(\alpha_j^0(y)|\, y) \quad j = 1, ..., J \tag{17}$$

Then by construction, the observed equilibrium sorting of households by income within and among neighborhoods for V_a and F_{ia} is observationally equivalent to the one given by V_0 and F_{i0}. By condition (ii) the implied joint distribution of income and quantity are also the same. Q.E.D.

The indirect utility function is not fully identified for three reasons. First, the number of neighborhoods is finite. Hence there are only a finite number of observations of p. Second, for those fixed values of p, the housing demand function $q(p, y)$ can only be traced out in the y dimension. Hence we can only identify J Engel curves for different levels of p. The dependence of q on p can only be bounded for values of p that are not observed in the data. Hence neither $q(\cdot)$ nor $V^b(\cdot)$ are fully identified. As the number of neighborhoods goes to infinity and market shares get arbitrarily small, these restrictions become unimportant. Third, $V^g(\cdot)$ is only identified from the conditions that state that the $J-1$ boundaries cannot intersect given the observed levels of p and g. As a consequence, the demand model considered above is only partially identified.

To obtain stronger results, we adopt a semi-parametric framework and introduce a parametrization of the indirect utility function while remaining as flexible as possible regarding the distribution of (α, y). For concreteness, we also adopt a form for $V^b(y,p)$ that implies constant price and income demand elasticities.

Assumption 6 *The utility function is known up to a finite vector of parameters, θ, and takes the form:*

$$V(\alpha, y, g_j, p_j) \;=\; \left\{\alpha\, g_j^{\rho} \;+\; \left[e^{\frac{y^{1-\nu}-1}{1-\nu}}\; e^{-\frac{Bp_j^{\eta+1}-1}{1+\eta}}\right]^{\rho}\right\}^{\frac{1}{\rho}} \tag{18}$$

where $\theta = (\rho, \eta, \nu, B)$ and $\rho < 0$, $\eta < 0$, $\nu > 0$, and $B > 0$.

Assumption 6 implies that the set of households that are indifferent between adjacent neighborhoods is characterized by the following equation:

$$\alpha_j(y) \;=\; \left[e^{\frac{y^{1-\nu}-1}{1-\nu}}\right]^{\rho}\; \frac{Q(p_j) - Q(p_{j-1})}{g_{j-1}^{\rho} - g_j^{\rho}} \tag{19}$$

where $Q(p_j) = e^{-\,\rho\,\frac{Bp_j^{\eta+1}-1}{1+\eta}}$ and the choice specific intercept is defined as:

$$K_j \equiv\; \ln\left(\frac{Q(p_j) - Q(p_{j-1})}{g_{j-1}^{\rho} - g_j^{\rho}}\right) \tag{20}$$

Roy's identity implies that the demand function is given by

$$q(p, y) \;=\; B\, p^{\eta}\, y^{\nu}. \tag{21}$$

Note that η is the price elasticity and ν is the income elasticity. B is the scale parameter of the demand equation. These three parameters are identified since they appear in the demand function.[19]

[19] To avoid stochastic singularities, we can easily extend the framework discussed above and assume that the housing demand or expenditures are subject to an idiosyncratic error that is revealed to households after they have chosen the neighborhood. This error term thus enters the housing demand, but does not affect the

Proposition 2 implies that ρ cannot be fully identified since this parameter does not appear in the demand function. The only restriction imposed on ρ is that the boundaries evaluated at ρ cannot intersect. Write the boundaries as $\alpha_j(y|\rho)$ to denote the dependence on ρ. Then vertical product differentiation implies that

$$\alpha_j(y|\rho_0) \quad < \quad \alpha_{j+1}(y|\rho_0) \quad j = 1, ..., J \tag{22}$$

As shown in the proof of Proposition 2, for any other ρ that is consistent with these inequality constraints, there exist a distribution of income and tastes that yields observationally equivalent sorting and household demand. The discussion in the previous section directly implies the following proposition:

Proposition 3 *The following results hold for our semiparametric model:*

1. *The three parameters of the demand equation* (ν_0, η_0, B_0) *are identified from the observed joint distribution of quantity and income given the price variation observed in the market.*

2. *For each household type* i *we can identify* $J-1$ *points of* $F_{i0}(\alpha|\, y)$ *if we know* ρ_0*.*

3. *The set of* ρ*'s that are consistent with the observed equilibrium outcomes is defined as:*

$$\left\{\rho \,\middle|\, \alpha_{j+1}(y|\rho) > \alpha_j(y|\rho) \quad \forall j\right\} \tag{23}$$

   This set contains ρ_0*.*

We can thus either identify or construct bounds for the parameters of interest.

---

neighborhood choice. An appendix that discussed this extension is available upon request from the authors. Alternatively, we can assume in estimation that observed housing demand is subject to measurement error. We follow that approach in our application.

4 Estimation

4.1 Measurement Error

The proofs of identification are constructive and can be used to devise a new estimator for this model.[20] For estimation purposes, it is desirable to relax assumption 5. The quality of a neighborhood may not be perfectly observed by the econometrician. In most applications, it is likely that we will observe some measures of quality. However, our observed measures may be subject to measurement error. We therefore make the following assumption:

Assumption 7 *We we do not observe g_j, but we observe, $\tilde{g}_j$, which is given by:*

$$\tilde{g}_j \quad = \quad g_j \; + \; \epsilon_j \tag{24}$$

where ϵ_j denotes measurement error.

Our model implies that there exists a monotonically increasing function which maps the price rank of a product into product quality. Let us denote the rank of neighborhood j by r_j. Hence the ascending bundles property implies that the following equation holds in equilibrium

$$g_j \quad = \quad g(r_j) \tag{25}$$

for some unknown monotonically increasing function $g(\cdot)$. Substituting equation (25) into equation (24), we obtain

$$\tilde{g}_j \quad = \quad g(r_j) \; + \; \epsilon_j \tag{26}$$

[20] As discussed in the previous section, we assume that we have consistent estimators of B, η and ν. As discussed in the next section estimation of these parameters is straight-forward and can be done prior to estimating the other parameters and functions of the model.

Furthermore suppose that $E[\epsilon_j|r_j] = 0$, i.e. the error term in equation (24) is conditionally independent of the rank of the neighborhood. In that case $g(r_j)$ is nonparametrically identified.[21] We estimate the function $g(r)$ using locally-linear kernel regression, as suggested by Fan (1992).[22]

Note that the estimation procedure can also be extended to account for multiple neighborhood characteristics as long as neighborhood quality satisfies an index assumption. Suppose we observe a vector of characteristics x_j. Assume that household preferences only depend on the linear index $g_j = x_j'\gamma + \epsilon$. We can modify the estimation procedure by combining the techniques discussed above with those suggested by Robinson (1988). Since quality can be measured on an arbitrary scale, we can normalize the coefficient of one the components in the index to be equal to plus or minus one to achieve identification. We implement this estimator in our application.

4.2 Imposing the Shape Restrictions

For the model to be well defined, we also need that the neighborhood specific intercepts are monotonically increasing: $K_1 < ... < K_J$. A necessary but not sufficient condition for this to hold is that $g(r)$ is monotonically increasing in r. If $g(r)$ has a sufficient degree of curvature, i.e. if the differences in neighborhood quality are sufficiently large relative to the differences in observed prices, then the intercepts are also monotonically increasing functions. It is therefore desirable to impose these shape restrictions in estimation. Moreover, by testing whether these curvature restrictions hold in the data, we can determine which values of ρ are admissible.

We impose these curvature restrictions using the isotone-kernel regression estimator proposed by Mammen (1991). This estimator uses a two step procedure to recover the monotonic function. First, a nonparametric estimator $\hat{g}(\cdot)$ is obtained, using local linear kernel regression. This function may not be monotonically increasing. In the second step,

[21] Instead of estimating g as function of rank, one estimate also estimate $\tilde{g}_j = g(p_j) + \epsilon_j$. In our application, we find that the two different approaches yield very similar results.

[22] For an overview of nonparametric techniques see, for example, Pagan and Ullah. (1999).

the estimated function is projected onto the space of shape-restricted (e.g. monotonic) functions. The estimator is defined by:

$$g^{sr}(r) = \text{argmin}_{g \in G} \int (g(r) - \hat{g}(r))^2 \, dr \tag{27}$$

where G denotes the class of shape-restricted functions.

We define $m = -g^{\rho}$ and $\hat{m} = \hat{g}^{\rho}$. The curvature restrictions then imply that $m(r)$ is sufficiently monotonically increasing in r. Our constrained estimator of the function is then obtained by minimizing the following objective function:

$$\min \sum_{j=1}^{J} (m_j - \hat{m}_j)^2 \tag{28}$$

subject to the constraints that:

$$\begin{aligned} m_{j-1} &< m_j - \delta_2 \\ \frac{1}{Q_{j-1} - Q_{j-2}}(m_{j-1} - m_{j-2}) &\leq \frac{1}{Q_j - Q_{j-1}}(m_j - m_j) - \delta_1 \end{aligned} \tag{29}$$

for non-stochastic constants $\delta_1, \delta_2 > 0$. Mammen (1991) shows that this estimator is consistent and derives rates of convergence.

We implement this estimator using quadratic programming techniques. Note that this estimator depends on ρ since the restrictions depend on ρ. Our procedure for recovering the identified set requires estimating the shape-restricted function on a grid of points for ρ.[23]

4.3 Testing the Shape Restrictions and Bounding ρ

Proposition 3 of the paper shows that there exists a set of ρ's that are consistent with the ascending bundles property. Once we allow for measurement error in quality, we need to

[23]Standard errors and confidence bands can be computed using bootstrap techniques. Bootstrap techniques are discussed in Efron and Tibshirani (1993) and Hall (1994).

test whether the shape restrictions are valid given a parameter ρ. In this section, we discuss how to test these shape restrictions and construct bounds for ρ. Our test procedure follows the basic ideas outlined in Hall and Yatchew (2005) (HY).

To develop our test, let us normalize the rank of a community to be between 0 and 1, i.e. $r_i = i/N$. Note that the rank is, by construction, uniformly distributed. Let $A \subset [0,1]$ denote the area of integration. For example, $A = [0,a] \cup [b,1]$ with $a < b$ would cover the "tail" regions where we expect violations of the null.

Consider the test statistic

$$T_A = \int_A \Big(g(r) - g^{sr}(r)\Big)^2 \, dr \tag{30}$$

where $g^{sr}(r)$ is the function in the shape restricted class that is closest to $g(r)$ under the L^2 norm.

A feasible test statistic is then given by:

$$\hat{T}_A = \sum_{i, r_i \in A} \Big(\hat{g}(r_i) - \hat{g}^{sr}(r_i)\Big)^2 \tag{31}$$

As pointed out by HY, we have strong reasons to believe that under the null $\sqrt{N}\hat{T}_A \to V$ where V is a continuous random variable. However, deriving the asymptotic distribution of the test statistic is difficult.

HY suggest employing the bootstrap to compute the critical values for this test statistic. To use the bootstrap procedure, we need to construct the distribution of $\hat{T}_A$ under the null hypothesis. We cannot sample from the underlying empirical distribution of the data since we do not know whether the empirical distribution satisfies the null hypothesis. Hence, we need to devise a bootstrap sampling algorithm which imposes the null on the data generating process.

To see how that can be done, consider the shape restricted estimator and define the residuals of the restricted model to be:

$$\hat{\epsilon}_i^{sr} \;=\; Y_i \;-\; \hat{g}^{sr}(r_i) \tag{32}$$

We recenter the residuals by subtracting the sample mean of the residuals from the individual residuals. The bootstrap then samples from the recentered distribution of the $\hat{\epsilon}_i^{sr}$. Let $\epsilon_1^s, ...\epsilon_N^s$ denote a bootstrap sample and define

$$Y_i^s \;=\; \hat{g}^{sr}(i/N) \;+\; \epsilon_i^s \tag{33}$$

For each bootstrap sample s we then compute the unconstrained and the constrained estimator and thus evaluate the test statistic $\hat{T}_A^s$.

By repeating this procedure S times, we can trace out the distribution of the test statistic under the null hypothesis

$$Pr\{\hat{T}_A \leq x\} \;\approx\; \frac{1}{S}\sum_{s=1}^{S} 1\{\hat{T}_A^s \leq x\} \tag{34}$$

It is straight-forward to read off the critical values of our test statistic from the distribution above. We compare the realization of our test statistic in equation (31) with the critical value and determine whether or not to reject the null.

To determine which $\rho' s$ are feasible, we pick a grid of values for ρ. We then implement the test procedure for each value of ρ and determine the set of ρ's that are consistent with the ascending bundles property.

4.4 Estimating the Conditional Distribution of Tastes

The constrained estimator of $g(\cdot)$ directly implies an estimator of the boundary indifference loci $\alpha_j(y|\rho)$ which are well behaved. We denote these estimators by $\hat{\alpha}_j(y|\rho)$. Note that the constrained estimator of the function $g(\cdot)$ depends on ρ since the constraints are functions of ρ. As a consequence $\hat{\alpha}_j(y|\rho)$ also depends on ρ.

We observe the empirical market shares and income distributions for each neighborhood, $\{n_{ij}^N, f_{ij}^N(y)\}_{j=1}^J$, where N denotes the relevant sample size. Following the discussion of identification in the previous section, a nonparametric estimator of the $J-1$ points of the conditional distribution of tastes given income, $\hat{F}_i^N(\alpha|\, y)$, is then given by:

$$\hat{F}_i^N(\hat{\alpha}_j(y|\rho)|\, y) \quad = \quad \sum_{k=1}^{j} \frac{f_{ik}^N(y)}{f_i^N(y)}\, n_{ik}^N \quad j = 1, ..., J-1 \tag{35}$$

where $f_i^N(y)$ denotes density that corresponds to the empirical income distribution of type i households in the market. It is straightforward to show that for any j

$$\sqrt{hN}\,\left(\sum_{k=1}^{j} \frac{f_{ik}^N(y)\, n_{ik}^N}{f_i^N(y)} \; - \; \sum_{k=1}^{j} \frac{f_{ik}(y)\, n_{ik}}{f_i(y)}\right) \xrightarrow{d} N(0, \sigma_j^2(y)) \tag{36}$$

where the asymptotic variance $\sigma_j^2(y)$ can be computed easily from the variances of the density estimators using the delta-method. This estimator depends on ρ. Thus for any admissible value of ρ we obtain a different conditional distribution of tastes.

5 Data

Our application focuses on residential choices and housing demand in Allegheny County, which includes Pittsburgh as its central city. Allegheny County consists of about 130 municipalities. Since the City of Pittsburgh is large, both in land area and population, we divide Pittsburgh based on its 32 wards. This leaves us with a total of 150 communities.

To measure the quality of public good provision, we employ measures of education quality, crime, and travel time to the city center. We construct an education index based on the PSSA, a math and reading test administered in all public schools for grades 5, 8, and 11 in Pennsylvania in the school year 1999-2000. Data on participation rates and average scores are available for each of the six tests for school districts and individual schools. There are no missing observations and participation rates are high. We average the six scores, weighted by enrollment in the different grades. For municipalities outside of Pittsburgh, a single school district sometimes serves several municipalities. We assign each of these municipalities the score for the school district. Getting education scores for the wards within the city of Pittsburgh is considerably more difficult. Here we rely on data reported by individual schools. School attendance zones for elementary schools, middle schools, and high schools sometimes overlap with the boundaries of the wards. In these cases we average the scores of all schools serving a ward weighted by the fraction of households served by that school.

To construct a crime index, we rely on two data sources- the Uniform Crime Report from the years 1990, 1999, 2000, and 2001 and data collected by the Pittsburgh Post-Gazette in 2001. The Uniform Crime Report is a yearly survey of the number and types of crime in each municipality in the U.S. It reports the number of actual incidents as reported by the police for murder, rape, robbery, assault, burglary, larceny, theft, as well as other crimes. We also construct a crime index for each of the 32 wards within the city of Pittsburgh based on data reported by the Pittsburgh Post-Gazette in 2001. We adjust these numbers by a multiplicative factor such that the numbers reported by the Post-Gazette for Pittsburgh as a whole match the numbers in the Uniform Crime Report.

The rush hour travel time to the central city of each municipality outside of Pittsburgh is taken from a data set provided by the Southwestern Pennsylvania Commission.

Demographic, income, housing and rental data are based on the U.S. Census. To compute a property tax rate in a community, we divide the aggregate housing value by the aggregate property taxes paid in each community.

We have obtained a detailed data set on local housing markets in Allegheny County. This data set contains housing prices and housing characteristics for essentially all residential properties in Allegheny County. Our data set consists of 93,763 properties which were recently sold. The data set contains a detailed list of housing characteristics including grade and condition assigned by an assessor of the property, year built, type of residence, finished living area, total number of rooms, number of bedrooms, number of full bathrooms, number of half bathrooms, whether the residence has a fireplace, and whether the residence has central air conditioning.

Housing values, v_{jn} are converted into imputed rents, r_{jn}, using the formula suggested by Poterba (1992):

$$r_{jn} = [(1 - \tau_y)(i + \tau_p) + \beta + m + \delta - \pi]\ v_{jn} \tag{37}$$

with average income tax rate $\tau_y = 0.15$, nominal interest rate $i = 0.079$, risk premium $\beta = 0.04$, maintenance minus depreciation $m - \delta = 0.02$, and inflation rate $\pi = .0286$.

Some communities in Allegheny County also rely on local income taxes. We convert local income taxes into implied property tax rates. Let τ_j^p be the property tax rate and τ_j^y be the local income tax rate in community j. We compute a property tax rate τ_j that yields the same revenue from the mean household in each community:

$$\tau_j = \tau_j^p + \tau_j^y \tfrac{\bar{y}_j}{\bar{v}_j} \tag{38}$$

To obtain housing prices, we estimate a hedonic regression of the form,

$$\ln r_{jn} = \sum_{j=1}^{J} I_j \ln p_j + \delta z_n + \epsilon_n \tag{39}$$

where r_{jn} is the imputed rent of house n in community j. The community specific intercepts of a regression with fixed effects can be interpreted as housing price estimates as discussed in detail in Sieg, Smith, Banzhaf, and Walsh (2002). The R^2 for the housing price regression is 0.5. Hence we control for much of the differences in the quality of housing across communities. Housing prices after taxes range from 1.06 to 6.96. Summary statistics of the sample of communities are reported in Table 1.

After estimating housing prices, we estimate the parameters of the housing demand equation using aggregate Census data. Our model implies that

$$\ln r_{jq} - \ln p_j = \ln \beta + \nu \ln y_q + \eta \ln p_j + \epsilon_{jq}^h \tag{40}$$

Table 1: Descriptive Statistics

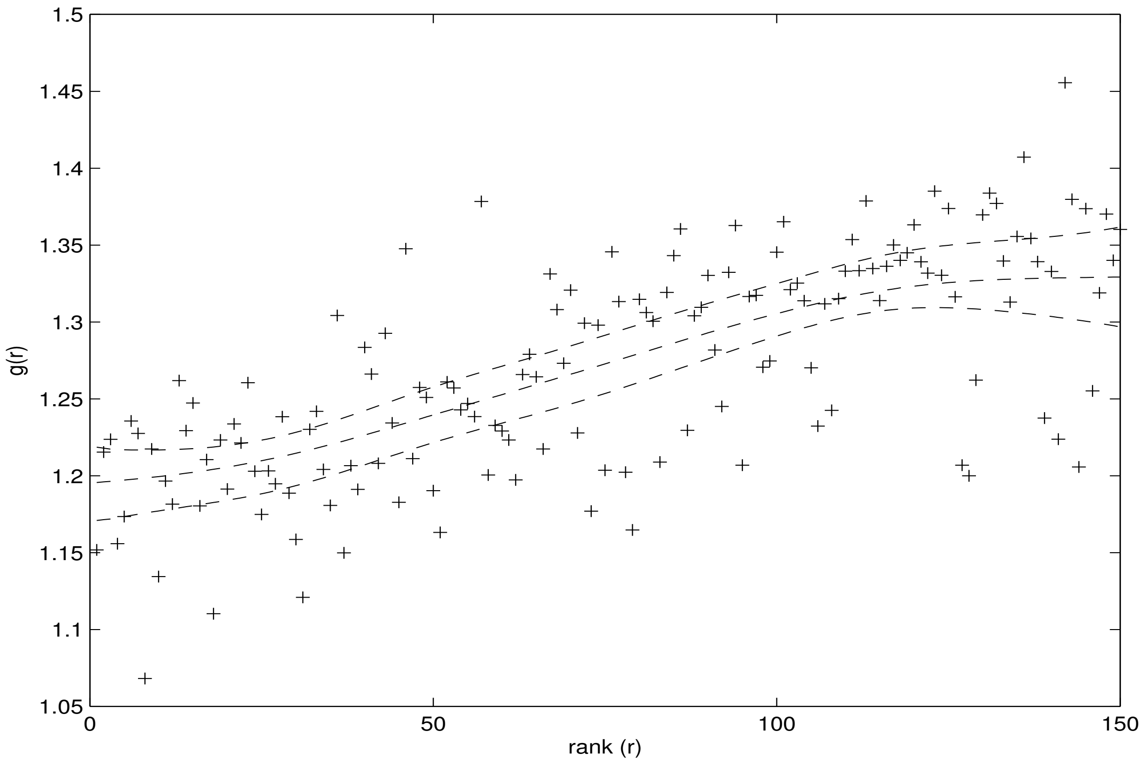
Variable	Mean	Std Dev	Minimum	Maximum
Population	8539	8551	467	46809
Number of Households	3581	3538	204	19467
Percent with Children	0.2665	0.0718	0.0632	0.5145
Price (before taxes)	2.9527	0.9403	1.0000	6.7302
Price (after taxes)	3.0919	0.9685	1.0638	6.9637
Education Index	1.2944	0.0836	1.0917	1.4699
Total Crime Index	690	888	0	8197
Property Tax Rate	0.0202	0.0026	0.0140	0.0283
Rush hour travel time	24.22	11.15	1	57
Income Tax Rate	0.0426	0.0077	0.0380	0.0568
Imputed Total Tax Rate	0.0490	0.0094	0.0319	0.0784
Mean Income	52947	30350	19580	233674
Mean Housing Value	100250	72587	26658	519080
Mean Rent	414	149	209	1156

where ϵ_{jq}^{h} denotes the error term in the model which may be due to measurement error in housing expenditures. r_{jq} represents the qth housing quantile and y_{jq} represents the respective income quantile. We use the 10% though 90% deciles in our estimation. The estimation results for the housing demand model yield an estimate for ν of 0.784 (0.017). The price elasticity η is estimated at -0.514 (0.027) and the intercept B is equal to 1.161 (0.182). The adjusted R^2 of the regression is 0.877.

6 Empirical Results

We estimate the function $g(r)$ using locally linear kernel regressions. The results of the estimation are plotted in Figure 2. We find that the smooth estimator $\hat{g}(r)$ does not violate the monotonicity condition. However, the unconstrained estimate of the function does not have enough curvature to ensure that the neighborhood specific intercepts are increasing.[24]

Figure 1: Estimation of $g(r)$



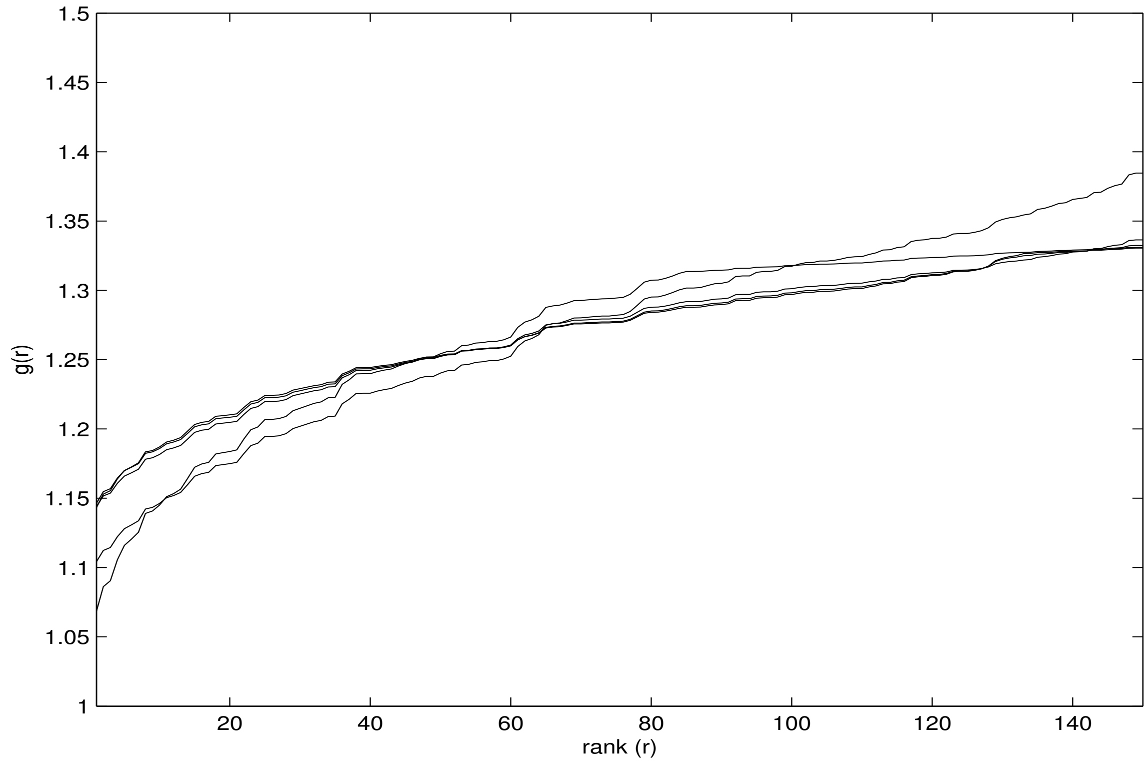
Notation: — unconstrained function, - - confidence bounds, + data.

We, therefore, impose the curvature restriction implied by the vertical differentiation property and implement the restricted estimator of the $g(\cdot)$ function. Since the constraints depend on the value of ρ, we estimate a number of constrained functions using values of ρ ranging from -0.1 to -1.1.[25] The results of these different estimators are plotted in Figure 3.

[24]We also control for differences in crime and commuting time to the city center in estimation using a partially linear estimator suggested by Robinson (1988). The point estimate of the coefficient of crime is 0.0059 (0.0034) and the coefficient for travel time is -0.0024 (0.0005).

[25]Using the Epple-Sieg (1999) parametric framework, we obtain a point estimate of -.198.

Figure 2: Constrained Estimation of $g(r)$

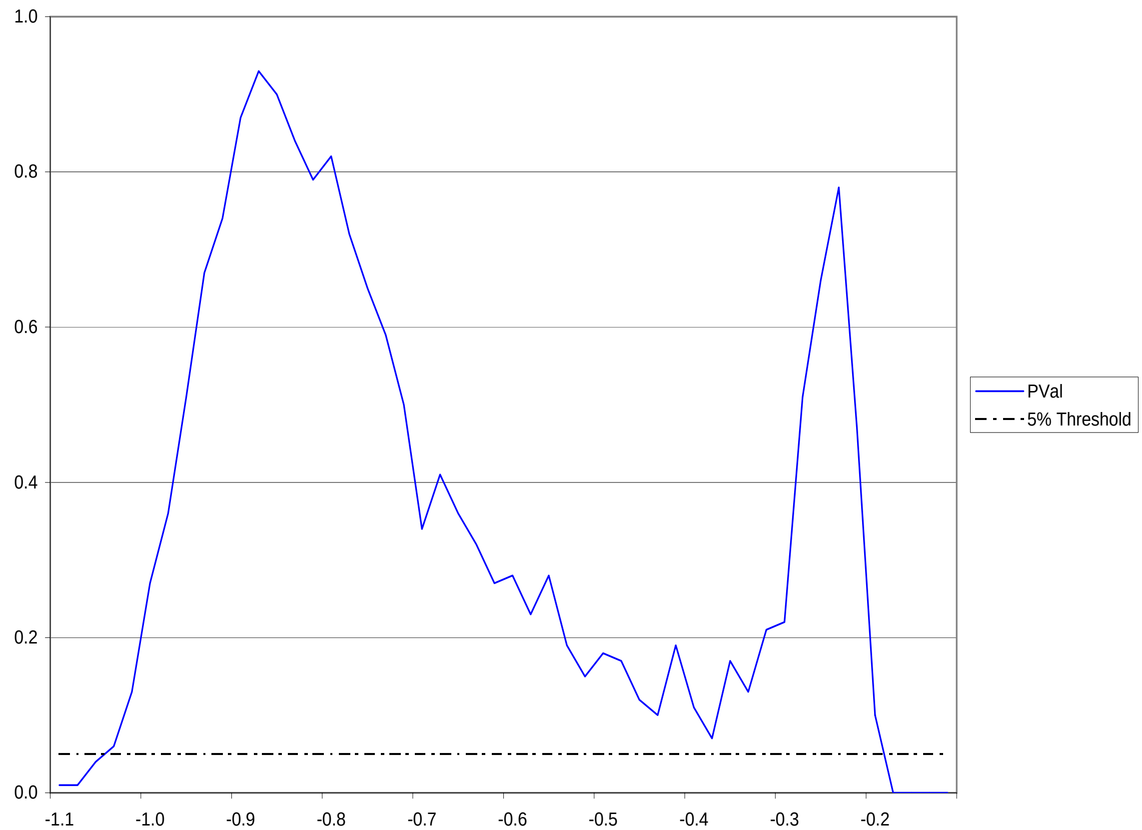


The figure plots the constrained function for ρ ranging between -0.1 and -0.9.

We find that the constrained functions have similar shapes for values of ρ ranging approximately from -.2 to -.8. As one goes outside this interval, we find that we need more curvature, especially in the tails of the function to accommodate the shape restrictions.

Next we implement our testing procedure to determine the admissible set for ρ as described in detail in in Section 4.3. For each value of ρ we plot the p-value associated with the HY test statistic in Figure 3. The HY test works fairly well in our application. Using a 95 percent level test, we find that the admissible set of ρ is given by the interval $[-0.94, -0.10]$. These results are qualitatively and quantitatively consistent with the graphical analysis in Figures 1 and 2. Once we move outside this interval, the restricted function significantly differs from the unrestricted function, which suggests that model violates the

Figure 3: Testing for Admissible Values of ρ



shape restrictions imposed by the ascending bundles property.[26]

Given the admissible set of ρ's, we then estimate the conditional distribution of tastes for quality. In our application we observe the sorting by households across communities. In our sample, 26.7 % of the households living in Allegheny county have children. The average income of these households is $66858 with a standard deviation of $67655. Households without children have an average income of $47,803 with a standard deviation of $53056. To characterize the observed sorting of household types across communities, we compute the following cumulative probabilities:

$$\sum_{k=1}^{j} \frac{f_{ik}^N(y)}{f_i^N(y)} \, n_{ik}^N$$

Recall that these probabilities measure the share of type i households with income level y that live in communities which have housing prices less than or equal to the price of community j.

Table 2 report these cumulative probabilities for households with and without children. We compute these probabilities for four different income levels: $19,330, $37921, $66,247 and $102,239. These income levels correspond to the 25th, 50th, 75th, and 90th percentile of the income distribution in the metropolitan area. We compute these measures for each decile of the housing price distribution. Each decile corresponds to 15 communities in our sample.

Table 2 provides some interesting new insights. Consider low income households with annual income of $19,330. 23 % of households with children live in the 15 cheapest communities, i.e. lowest decile of communities. Only 11 % of households without children live in these communities. The opposite pattern holds for richer households with annual income of $102,239. 57 % of households with children live in the 45 most expensive communities. Here the corresponding number for households without children is 52 %. We thus conclude that the sorting of households with children exhibits more stratification by income than

[26] We also implemented a second test that is based on an equally spaced grid of points. The alternative procedure produced qualitatively similar results, and indicated that the identified set is $[-0.78, -0.16]$.

Table 2: Household Sorting

Households With Children				
housing price	income			
percentile	19,330	37,921	66,247	102,239
10	0.231660	0.090155	0.044402	0.014032
20	0.339434	0.154663	0.080393	0.035698
30	0.467073	0.258031	0.148607	0.060395
40	0.627718	0.447927	0.285855	0.137517
50	0.704364	0.545725	0.374877	0.218680
60	0.779916	0.655959	0.493883	0.353727
70	0.837322	0.743523	0.586674	0.432869
80	0.900829	0.862824	0.776079	0.622362
90	0.958392	0.956477	0.933042	0.857095
Households Without Children				
housing price	income			
percentile	19,330	37,921	66,247	102,239
10	0.109794	0.058770	0.051086	0.029720
20	0.194341	0.127941	0.100536	0.064322
30	0.292824	0.215270	0.176346	0.099569
40	0.451452	0.372183	0.327224	0.206174
50	0.517419	0.453446	0.414252	0.276741
60	0.641752	0.571688	0.527462	0.398636
70	0.710403	0.655799	0.612347	0.480474
80	0.826984	0.790920	0.763523	0.649749
90	0.912979	0.917747	0.899911	0.834027

the observed sorting of households without children. Households with children and income levels below the median 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 tastes for high price (and high amenity) communities than households without children.

For any value of ρ that is admissible, we then compute the boundaries between adjacent communities. Given the parametrization of our utility function, these boundaries are given by the following expression:

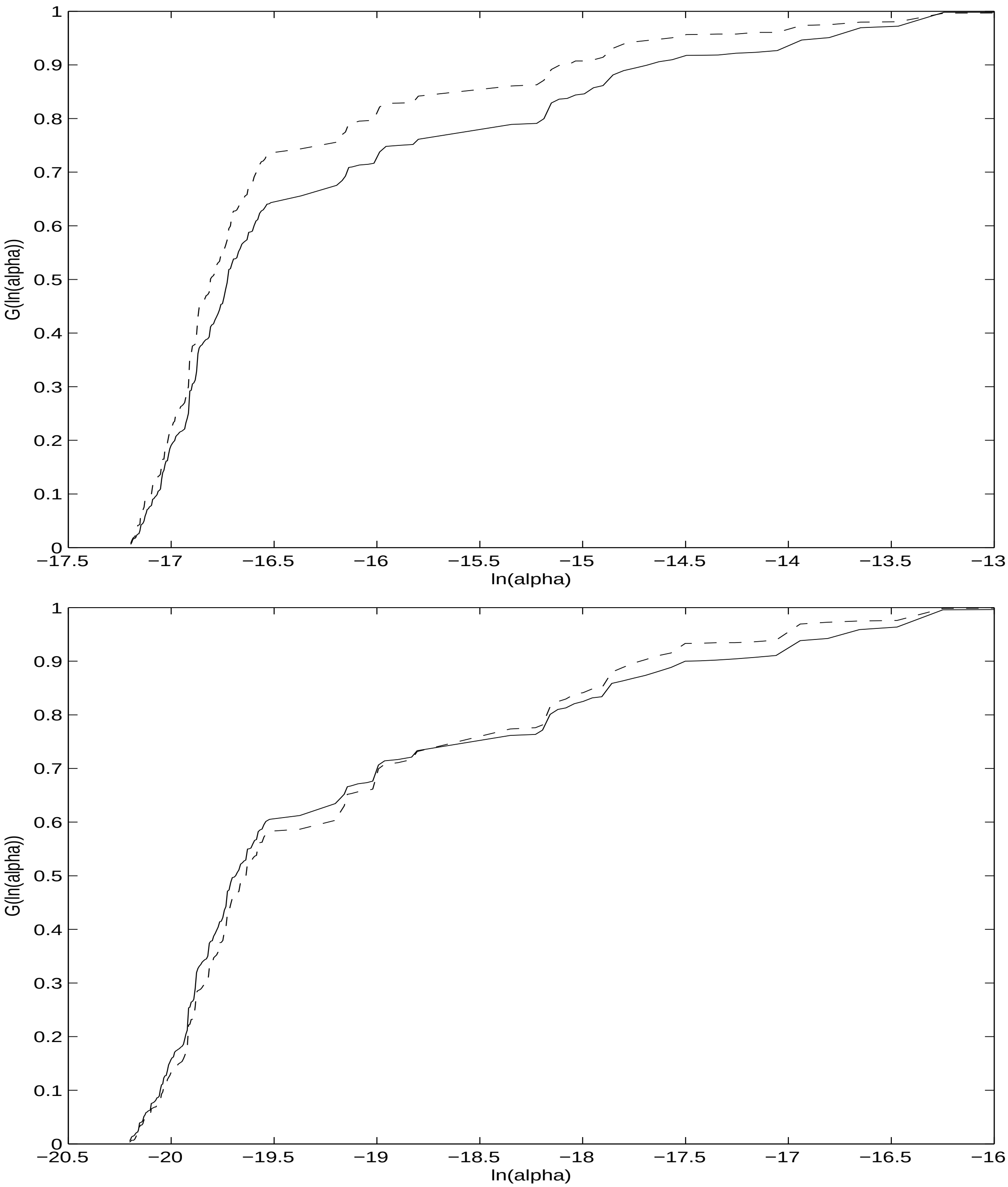
$$\ln(\alpha_j) \;=\; K_j \,+\, \rho\,\frac{y^{1-\nu}-1}{1-\nu} \tag{41}$$

Given these boundaries and the results in Proposition 3, it is then straight-forward to construct the conditional distribution of tastes given income. To illustrate the algorithm, we set $\rho = 0.52$ which is the mid point of the admissible interval of ρ based on the HT test. In Figure 4 plot the distribution for tastes conditional on two income levels ($37921 and $66,247). The solid (dashed) line denotes households without (with) children. We find that higher income households have on average significantly lower tastes for local public goods than lower income households. This result is true for households with and without children. Lower income families with children tend to have lower tastes for public goods than households without children. We also computed the conditional distribution of tastes for other feasible values of ρ and the results were qualitatively the same.

7 Conclusions

We have discussed nonparametric identification of models of locational equilibrium. Our results show that the model considered in this paper is partially nonparametrically identified. The proofs of (partial) identification are constructive. We have shown how to derive a new two-step estimator for the semiparametric model. This estimator differs significantly from previously used parametric estimators which are typically based on share inversion al-

Figure 4: Conditional Distribution of Tastes



The top (bottom) panel shows distribution conditional on income equal \$37921 (\$66,247). The solid (dashed) line denotes households without (with) children.

gorithms. We have discussed the asymptotic properties of the new estimator and provided some simple algorithms which can be used to implement the estimator.

We have studied residential sorting and housing demand in a new application that focuses on sorting of households with and without children across municipalities in Allegheny County. We have documented the observed sorting of each household type by income among the set of communities in our sample. Our empirical findings suggest that there are significant differences in the observed sorting patterns of household types. We find that households with children seem to be more sensitive to differences in housing prices and local public goods than households without children. The sorting pattern of households with children exhibit a lot more stratification by income than the corresponding pattern for households without children. Moreover, we find that low income households with children have on average lower tastes for public goods than households without children. The opposite is true for households with higher income levels.

We view the findings of this paper as encouraging for further research in this area. There seems to be ample scope for using non- and semiparametric estimation techniques to estimate richer specification of the type of locational equilibrium models considered in this paper. We have limited our discussion to hierarchical models in which there exists a clear ranking among the set of communities. If households have heterogeneous tastes defined over a vector of local public goods and amenities, household-sorting equilibria do not necessarily satisfy the ascending bundles property. Moreover, there will be both vertical and horizontal product differentiation in equilibrium. Alternatively, one could consider extensions of the hierarchical model which allow for differences in utility functions across types or additional unobserved heterogeneity in tastes for housing. For low income households with children providing basic necessities such as food and shelter may take precedence over concerns for local public goods. Establishing conditions for non- or semiparametric identification and developing feasible estimators for these models is an important area for future research.

References

Bajari, P. and Benkard, L. (2005). Demand Estimation With Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach. *Journal of Political Economy*, *113*, 1239–77.

Bajari, P. and Kahn, M. (2004). Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities. *Journal of Business and Economic Statistics*, *23(1)*, 20–33.

Bayer, P., Ferreira, F., and McMillan, R. (2007). A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy*, *115 (4)*, 588–638.

Bayer, P., McMillan, R., and Reuben, K. (2004). The Causes and Consequences of Residential Segregation: An Equilibrium Analysis of Neighborhood Sorting. Working Paper.

Berry, S. (1994). Estimating Discrete-Choice Models of Product Differentiation. *Rand Journal of Economics*, *25(2)*, 242–262.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, *63 (4)*, 841–890.

Blundell, R., Browning, M., and Crawford., I. (2003). Nonparametric Engel Curves and Revealed Preferences. *Econometrica*, *71*, 205–240.

Calabrese, S., Epple, D., Romer, T., and Sieg, H. (2006). Local Public Good Provision: Voting, Peer Effects, and Mobility. *Journal of Public Economics*, *90 (6-7)*, 959–81.

Chernozhukov, V., Imbens, G., and Newey, W. (2006). Nonparametric Identification and Estimation of Non-separable Models. *Journal of Econometrics*, *forthcoming*.

Efron, B. and Tibshirani, R. (1993). *Introduction to the Bootstrap*. Chapman and Hall.

Ekeland, I., Heckman, J., and Nesheim, L. (2004). Identification and Estimation of Hedonic Models. *Journal of Political Economy*, *112(1)*, S60–S101.

Epple, D. and Romer, T. (1991). Mobility and Redistribution. *Journal of Political Economy*, *99(4)*, 828–858.

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

Epple, D. and Sieg, H. (1999). Estimating Equilibrium Models of Local Jurisdictions. *Journal of Political Economy*, *107 (4)*, 645–681.

Fan, J. (1992). Design-adaptive Nonparametric Regression. *Journal of the American Statistical Association*, *87*, 998–1004.

Fernandez, R. and Rogerson, R. (1998). Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform. *American Economic Review*, *88(4)*, 813–833.

Ferreira, F. (2005). You Can Take It with You: Proposition 13 Tax Benefits, Residential Mobility, and Willingness to Pay for Housing Amenities. Working Paper.

Ferreyra, M. (2007). Estimating the Effects of Private School Vouchers in Multi-District Economies. *American Economic Review*, *97*, 789–817.

Hall, P. (1994). Methodology and Theory for the Bootstrap. In *Handbook of Econometrics IV*. Elsevier North Holland.

Hall, P. and Yatchew, A. (2005). Unified approach to testing functional hypotheses in semiparametric contexts. *Journal of Econometrics*, *127*, 225–52.

Han, A. (1987). A Nonparamtric Analysis of Transformations. *Journal of Econometrics*, *35*, 191–209.

Heckman, J. and Honore, B. (1990). The Empirical Content of the Roy Model. *Econometrica*, *58 (5)*, 1121–49.

Heckman, J., Matzkin, R., and Nesheim, L. (2004). *Simulation and Estimation of Hedonic Models.* Frontiers of Applied General Equilibrium Modeling: Essays in Honour of Herbert Scarf. North Holland.

Imbens, G. and Manski, C. (2004). Confidence Intervals for Partially Identified Parameters. *Econometrica*, *72*, 1845–1857.

Kahn, S. (2001). Two-stage rank estimation of quantile index models. *Journal of Econometrics*, *100*, 319–55.

Lewbel, A. and Linton, O. (2003). Nonparametric Estimation of Homothetic and Homothetically Separable Functions. Working Paper.

Mammen, E. (1991). Nonparametric Regression under Qualitative Smoothness Assumptions. *Annals of Statistics*, *19*, 741–759.

Manski, C. (1997). Monotone Treatment Response. *Econometrica*, *65*, 1311–1334.

Matzkin, R. (1994). Restrictions of Economic Theory in Nonparametric Methods. In *Handbook of Econometrics IV*. Elsevier North Holland.

Nechyba, T. (1997). Local Property and State Income Taxes: the Role of Interjurisdictional Competition and Collusion. *Journal of Political Economy*, *105 (2)*, 351–84.

Nechyba, T. (2000). Mobility, Targeting and Private School Vouchers. *American Economic Review*, *90*, 130–146.

Nechyba, T. (2003). Centralization, Fiscal Federalism, and Private School Attendance. *International Economic Review*, *44*, 179–204.

Nesheim, L. (2001). Equilibrium Sorting of Heterogeneous Consumers Across Locations: Theory and Empirical Implications. Dissertation, University of Chicago.

Pagan, A. and Ullah., A. (1999). *Nonparametric Econometrics.* Cambridge University Press.

Poterba, J. (1992). Taxation and Housing: Old Questions, New Answers. *American Economic Review*, *82 (2)*, 237–242.

Powell, J. (1994). Estimation of Semiparametric Models. In *Handbook of Econometrics IV*. Elsevier North Holland.

Robinson, P. (1988). Root-N-Consistent Semiparametric Regression. *Econometrica*, *56*, 931–954.

Rothstein, J. (2006). Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition Among Jurisdictions. *American Economic Review*, *96*, 1333–1350.

Schmidheiny, K. (2006). Income segregation and local progressive taxation: Empirical evidence from Switzerland. *Journal of Public Economics*, *90*, 429–458.

Sieg, H., Smith, V. K., Banzhaf, S., and Walsh, R. (2002). Interjurisdictional Housing Prices in Locational Equilibrium. *Journal of Urban Economics*, *52*, 131–153.

Sieg, H., Smith, V. K., Banzhaf, S., and Walsh, R. (2004). Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods. *International Economic Review*, *45 (4)*, 1047–1077.

Tamer, E. (2003). Incomplete Simulataneous Discrete Choice Models with Multiple Equilibria. *Review of Economic Studies*, *70 (1)*, 147–166.

Walsh, R. (2007). Endogenous open space amenities in a locational equilibrium. *Journal of Urban Economics*, *61 (2)*, 319–344.

Wu, J. and S. (2003). Estimating households' preferences for environmental amenities using equilibrium models of local jurisdictions. *Scottish Journal of Political Economy*, *50*, 189–206.