Identification and Semiparametric Estimation of Equilibrium Models of Local Jurisdictions*

Dennis Epple
Carnegie Mellon University and NBER

Michael Peress
Carnegie Mellon University

Holger Sieg
Carnegie Mellon University and NBER

February 23, 2005

*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, Lanier Benkhard, Steve Berry, Martin Browning, Jan Bruckner, Andrew Chesher, Phil Haile, Lars Nesheim, Charles Manski, Rosa Matzkin, Ariel Pakes, Richard Romano, John Rust, Matt Shum, Chris Taber, 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 workshops at the University of California in Berkeley, Chicago, Florida, Johns Hopkins, Northwestern and Yale for comments and suggestions. Financial support for this research is provided by the NSF, the MacArthur Foundation, and the Alfred P. Sloan Foundation.

Abstract

Research over the past several years has led to development of models characterizing equilibrium in a system of local jurisdictions. More recently, there have been a number of studies which have estimated these models. One potential drawback of the approach adopted in these empirical studies is that identification of the parameters of the model relies on functional form assumptions on the distribution of unobserved tastes in the population. In this paper, we provide an analysis of identification and estimation of locational equilibrium models in non- and semiparametric frameworks. We show that a broad class of models is identified without imposing strong parametric restrictions on the distribution of unobserved tastes for local public goods. The proofs of identification are constructive and can be used to derive a new class of semiparametric estimators for these models. Our empirical results show that these estimators perform well in an application.

JEL classification: C51, H31, R12

1 Introduction

Research over the past several years has led to development of models characterizing equilibrium in a system of local jurisdictions. More recently, there have been a number of studies which have estimated these models. The evidence suggests that simple parametric models can explain the observed sorting of households by income among local jurisdictions reasonably well.[1] However the approach adopted in all previous empirical studies relies on parametric assumptions. The purpose of this paper is to provide a discussion of identification and estimation of equilibrium models of local jurisdictions within a framework which does not rely heavily on functional form assumptions.[2]

The starting point of our analysis is the locational equilibrium model considered in Epple and Sieg (1999). In our baseline model, heterogeneity among households is characterized by the joint distribution of income and tastes for public goods. We extend this models and allow for additional sources of observed heterogeneity among households. For example, it is reasonable to assume that households with children have a different distribution of tastes and income than households without children. We use simple mixtures of distributions to characterize heterogeneity in the extended model. 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 previous papers.

We study identification in a single cross-section of communities which form a metropolitan area. Since tastes for local public goods are inherently unobservable, it may be difficult to justify specific functional form assumptions imposed on the (conditional) distribution of tastes. We therefore derive conditions that allow us to nonparametrically identify the distribution of household types and the indirect utility function of households based on the

[1] See, for example, Epple and Sieg (1999), Epple, Romer, and Sieg (2001), Sieg, Smith, Banzhaf, and Walsh (2004), and Walsh (2002). See also Nesheim (2001), Ferreyra (2003), and Bayer, McMillan, and Reuben (2003) for related empirical approaches

[2] This research is thus similar to recent work by Ekeland, Heckman, and Nesheim (2004), Heckman, Matzkin, and Nesheim (2004) and Bajari and Benkard (2002) on identification and estimation of hedonic models.

observed equilibrium outcomes.[3]

We find that it is possible to nonparametrically identify a finite number of points of the distribution of tastes conditional on income for each discrete household type, if we know the indirect utility function of the households. These points correspond to the points on the boundary between adjacent communities. For points that are not on the boundary loci we can only provide lower and upper bounds for the distribution. These bounds tend to become tighter as the number of communities in the application increases. The discreteness of the choice set thus imposes limits to identification for the most general versions of our model.[4]

Joint nonparametric identification of the distribution of household types and the indirect utility function is more difficult to establish. We consider the case in which the utility function is separable in private and public good consumption. Separability implies that we can recover the sub-utility function which models preferences over private goods from the observed sorting by housing expenditures conditional on income among the set of communities. Moreover, we can provide nonparametric bounds for the sub-utility function which characterizes preferences over local public goods. We thus conclude that the model is partially nonparametrically identified.[5]

To obtain stronger identification results, we adopt a semiparametric approach and impose parametric assumptions on the function characterizing household preferences.[6] However, we do not impose any functional form assumptions on the joint distribution of tastes and income among households. A sufficient condition for point identification of the distribution of household types and the parameters of the indirect utility function is that there exists one discrete household type for which tastes for local public goods and income

[3]Our work is also closely related to Blundell, Browning, and Crawford. (2003) who discuss nonparametric tests of revealed preference models. Chesher (2003) considers nonparametric identification of derivatives of regression functions which vary across households that have identical covariates. Athey and Haile (2002) provide a general discussion of identification in auction models.

[4]Point identification cannot be achieved in many econometric applications. In that case attention naturally shifts to characterizing informative bounds parameters. Some recent examples are Manski (1997) and Tamer (2003).

[5]For a general discussion of partial identification see Imbens and Manski (2004).

[6]Powell (1994) provides an overview of semiparametric estimation.

are independently distributed. Independence of income and tastes is a strong assumption. However, the strategy used to prove identification in this case can be generalized to derive informative bounds for the parameters or functions of interest, or conditions for point identification, for more interesting models. We have argued above that it is reasonable to assume that households with children have a different distribution of tastes and income than households without children. The sorting of households along these observed dimensions provides additional information which helps to identify the underlying structure of the model. We can achieve point identification of the model if we impose some reasonable shape restrictions on the joint distribution of income and tastes for public goods for these discrete household types. This approach allows identification of the model even if the distribution of income and tastes does not satisfy an independence assumption for any discrete type.

While most of our analysis focuses on identification based on observing an equilibrium in one metropolitan area, we also briefly consider identification of the model if equilibria are observed in multiple markets. In principle, one can observe equilibria of different metropolitan areas at a single point of time or equilibria of the same metropolitan area at different points of time. Observing multiple equilibria alone is not a sufficient condition for identification. We must also impose cross-market restrictions on the distributions of income and tastes and the shape of preferences.[7]

Our proofs of identification are constructive. They give rise to estimation algorithms that can be used to estimate the parameters and functions of interest. We propose a new two-step semiparametric estimator for the model. Locational equilibrium implies that public good provision should be monotonically increasing in the price rank of a community. Moreover, this function must also have a sufficient degree of curvature to guarantee that the differences in public good provision between adjacent communities are large enough given the observed differences in housing prices. In the first step, we nonparametrically estimate a function which links public good provision to the observed price rank of a community using recent innovations in nonparametric estimation which impose monotonicity

[7]See also Epple (1987) for a discussion of identification of hedonic models using multiple markets.

and curvature constraints on the underlying function.[8] In the second step, we then develop a new (semiparametric) estimator of the joint distribution of tastes and income for each discrete type which is obtained by inverting the community specific income distributions. This estimator differs significantly from previously used parametric estimators. It does not rely on computationally intensive share inversion algorithms which are the cornerstone of many parametric estimators of differentiated product models.[9]

Finally, we study the properties of the estimators proposed in this paper in a new application. The empirical analysis is based on a new data set that we have assembled for the Pittsburgh metropolitan area. We find that there are significant differences in the observed sorting of households with and without children in our sample. In particular, sorting of households with children exhibits more stratification by income than sorting of 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. We find that our semiparametric estimator performs well in this application and significantly improves our understanding of observed household sorting patterns among a set of local jurisdictions.

The rest of the paper is organized as follows. Section 2 provides a review of locational equilibrium models. Section 3 discusses identification. Section 4 develops a new semiparametric two-step estimator which can be constructed based on the identification results. Section 5 presents our data and introduces the application studied in this paper. Section 6 reports the empirical findings. Section 7 offers some conclusions and discusses future research.

2 A Locational Equilibrium Model

In this section, we review the baseline equilibrium model and extend it to allow for additional sources of heterogeneity among households. This model considers the problem of

[8] Matzkin (1994) provides an overview of nonparametric estimators that impose shape restrictions.

[9] For a discussion of share inversion techniques see Hotz and Miller (1993) and Berry (1994).

public good provision and residential decisions in a system of multiple jurisdictions.[10] The economy consists of a finite number of communities and a continuum of households living in a metropolitan area. The homogeneous land in the metropolitan area is divided among a number of communities, each of which has fixed boundaries. Jurisdictions may differ in the amount of land contained within their boundaries.

We consider an economy with a finite number of types I that differ in their endowed income, y, and in a taste parameter, α, which reflects the household's strength of preferences for the public good. For example, it is reasonable to assume that households with children have a different distribution of income and tastes for public goods than households without children. Each household type i 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 with taste parameter α and income y is referred to as a tuple (α, y). A household living in a community has preferences defined over a local public good, g, a local housing good, h, and a composite private good, b.[11]

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

Denote with p_h the net-of-tax price of a unit of housing services in a community. Households pay taxes that are levied on the value of housing services. Let t be an *ad valorem* tax on

[10]The theoretical literature which provide the foundation for these papers was inspired by Tiebout (1956). See, for example, Epple, Filimon, and Romer (1984), Goodspeed (1989), Epple and Romer (1991), Nechyba (1997a, 1997b) and Fernandez and Rogerson (1996, 1998).

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

housing and hence the gross-of tax price of housing is given by $p = (1+t)p_h$. Households maximize their utility with respect to a budget constraint:

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

It is convenient to represent the preferences of a household living in community j using the indirect utility function, $V(\alpha, y, g_j, p_j)$, derived by solving the optimization problem given in equation (1). To characterize the equilibrium of 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 monotonically increasing in α.*

Let C_j denote the population living in community j:

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

The share of households of type i living in community 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 population share of community j:

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

Let $h(\cdot)$ denote the household housing demand function which can be derived using Roy's identity. The budget of community j must be balanced which implies that:

$$t_j \; p_j^h \; \left[\sum_{i=1}^{I} \; P_i \; \int_{C_j} \; h(p_j, \alpha, y) \;\; f_i(\alpha, y) \; dy \; d\alpha \right] \Big/ \; n_j \;\; = \;\; c(g_j) \tag{6}$$

where $c(g)$ is the cost per household of providing g. We assume that

Assumption 4 *$c(g)$ is an increasing and convex function that is twice differentiable in g.*

We assume that the pair (t, g) in each community is chosen by majority rule. In each community, voters take the (t, g) pairs in all other communities as given when making their decisions. One can make a variety of assumptions about voter sophistication regarding anticipation of the way changes in the community's own (t, g) pair affect the community's housing prices and migration into or out of the community. For example, utility-taking voters base their voting decisions on the housing price and migration effects that would occur if the utility in the next best alternative community is given. The community budget constraint, housing market clearing, and perceived migration effects define the function $p(g)$ that determines the government-services possibility frontier, i.e. GPF $= \{g(t), p(t) \; |t \in R^+\}$. For given tax and expenditure policies in other communities, a point on the GPF that cannot be beaten in a majority vote is a majority equilibrium.[12]

Mobility among communities is costless, and in equilibrium every household lives in his or her preferred community. To close the model we assume that the housing stock in each community is owned by absentee landlords. As a consequence, we can characterize housing supply in each community by a simple housing supply function, $H^s(p^h)$. Having specified all components of our equilibrium model, we define an intercommunity equilibrium as follows:

Definition 1 *An* **intercommunity equilibrium** *consists of a set of communities, $\{1, ..., J\}$; a continuum of households, C; a distribution, F_i, of household characteristics*

[12]Since our discussion of identification of the model does not depend on the way we specify the voting mechanism in each community, we do not present more details here.

α and y for each type i; and a partition of C across communities $\{C_1, ..., C_J\}$, such that every community has a positive population, i.e. $0 < n_j < 1$; a vector of prices and taxes, $(p_1^*, t_1^*, ..., p_J^*, t_J^*)$; an allocation of public goods, $(g_1^*, ..., g_J^*)$; and an allocation, (h^*, b^*), for every household (α, y), such that:

1. *Every household (α, y), living in community j maximizes its utility subject to the budget constraint:*

$$
\begin{aligned}
(h^*, b^*) \;&=\; \arg\max_{(h,b)} U(\alpha, g_j^*, h, b) \\
&\quad\; s.t. \;\; p_j^* \, h \;=\; y \;-\; b
\end{aligned}
$$

2. *Each household lives in one community and no household wants to move to a different community, i.e. for a household living in community j, the following holds:*

$$
V(\alpha, g_j^*, p_j^*, y) \;\geq\; \max_{i \neq j} V(\alpha, g_i^*, p_i^*, y) \tag{7}
$$

3. *The housing market clears in every community:*

$$
\left[\sum_{i=1}^{I} P_i \int_{C_j} h^*(p_j^*, y, \alpha) \; f_i(\alpha, y) \; dy \; d\alpha \right] \;=\; H_j^s\left(\frac{p_j^*}{1 + t_j^*}\right) \tag{8}
$$

4. *The budget of every community is balanced:*

$$
\frac{t_j^*}{1 + t_j^*} \, p_j^* \left[\sum_{i=1}^{I} P_i \int_{C_j} h^*(p_j^*, y, \alpha) \; f_i(\alpha, y) \; dy \; d\alpha \right] \Big/ \, n_j \;=\; c(g_j^*) \tag{9}
$$

5. *There is a voting equilibrium in each community: Over all levels of (g_j, t_j) that are perceived to be feasible allocations by the voters in community j, at least half of the voters prefer (g_j^*, t_j^*) over any other feasible (g_j, t_j).*

We assume that an equilibrium exists and study identification of the model based on an observed equilibrium in one metropolitan area.[13] The necessary conditions for locational equilibrium impose a number of restrictions on the equilibrium allocation that apply quite broadly, in that they do not depend on specific features of the model such as the collective choice mechanism that determines policy variables in each community, or the technology of producing the public good. Proposition 1 summarizes three necessary conditions that hold in equilibrium for communities that are not identical and, hence, differ in housing prices.

Proposition 1 *Consider an equilibrium allocation in which no two communities have the same housing prices. For such an allocation to be a locational equilibrium – no-one wishes to move – there must be an ordering of community pairs,* $\{(g_1, p_1), ..., (g_J, p_J)\}$*, such that:*

1. **Boundary Indifference:** *The set of "border" households between any two adjacent communities are indifferent between the two communities. This set is characterized by the following expression:*

$$R_j \;=\; \{(\alpha, y) \mid \; V(\alpha, g_j, p_j, y) \;=\; V(\alpha, g_{j+1}, p_{j+1}, y)\} \qquad j = 1, ..., J-1 \tag{10}$$

2. **Stratification:** *Let* $\alpha_j(y)$ *be the implicit function defined by equation (10). Then, for each level of income* y*, the residents of community* j *consist of those with tastes,* α*, given by:*

$$\alpha_{j-1}(y) \;<\; \alpha \;<\; \alpha_j(y) \tag{11}$$

3. **Increasing Bundles:** *Consider two communities* i *and* j *such that* $p_i \;>\; p_j$*. Then* $g_i > g_j$ *if and only if* $\alpha_i(y) \;>\; \alpha_j(y)$*.*

A formal proof of Proposition 1 is given in Epple and Sieg (1999). As we will see below, these necessary conditions of equilibrium are important in establishing identification of the model.

[13]For a rigorous discussion of existence see Calabrese, Epple, Romer, and Sieg (2004).

3 Identification

3.1 Nonparametric Identification

The nature of the identification problem is 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:

Assumption 5 *For every community in the metropolitan area the econometrician observes:*

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

as well as housing prices, p_j, tax rates t_j, and local public good provision g_j.

These types of data are available from the U.S. Census and state and local government publications. Hence we study identification based on publicly available data sources.[14]

First we consider the case in which $V(\cdot)$ is known to the econometrician.[15] 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 2 *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.*

[14] Assumption 5 also directly implies that the marginal distribution of income in the metropolitan area is observed by the econometrician.

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

Proof:

Note that the joint distribution of (α, y) of household type i in community j is given by:

$$f_{ij}(\alpha, y) \;=\; \begin{cases} \frac{f_i(\alpha,y)}{n_{ij}} & \text{if} \quad (\alpha, y) \in C_j \\ 0 & \text{if} \quad (\alpha, y) \notin C_j \end{cases} \tag{12}$$

Hence the income distribution of type i in community j is given by:

$$\begin{aligned} f_{ij}(y) &= \int_{C_j} f_{ij}(\alpha, y)\; d\alpha \\ &= \frac{f_i(y)}{n_{ij}} \int_{\alpha_{j-1}(y)}^{\alpha_j(y)} f_i(\alpha|y)\; d\alpha \\ &= \frac{f_i(y)}{n_{ij}} \; [F_i(\alpha_j(y)|\; y) - F_i(\alpha_{j-1}(y)|\; y)] \end{aligned} \tag{13}$$

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

$$F_i(\alpha_1(y)|\; y) \;=\; \frac{f_{i1}(y)}{f_i(y)}\; n_{i1} \tag{14}$$

and all other communities $j > 1$:

$$F_i(\alpha_j(y)|\; y) \;=\; \frac{\sum_{k=1}^{j} n_{ik} f_{ik}(y)}{f_i(y)} \tag{15}$$

We can, therefore, 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 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 communities in equilibrium, as long as the transformed distribution has the correct mass points at the boundaries. As a consequence the conditional distribution of tastes given income is not identified 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:

$$\underline{F}_i(\alpha|\ y) \ = \ F_i(\alpha_j(y)|\ y) \ \leq \ F_i(\alpha|\ y) \ \leq \ F_i(\alpha_{j+1}(y)|\ y) \ = \ \overline{F}_i(\alpha|\ y) \tag{16}$$

If there are many small communities in the metropolitan area, we would expect that the difference between $\alpha_j(y)$ and $\alpha_{j+1}(y)$ will be small for most adjacent communities.[16] 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 $\{F_i(\cdot)\}_{i=1}^{I}$. In order to obtain further results it is useful to impose additional structure on the problem. We consider, in the following discussion of identification, a preference function that is sufficiently general to subsume specifications that have generally been adopted in applied equilibrium models of local jurisdictions.

Assumption 6 *The indirect utility function is additively separable in the sub-utility function for the public good and the sub-utility for the private goods bundle and hence can be written as:*

$$V(\alpha, y, g, p) \ = \ \alpha \, V^g(g) \ + \ V^b(y, p) \tag{17}$$

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

$$h(p, y) \quad = \quad -\frac{\partial V^b / \partial p}{\partial V^b / \partial y} \tag{18}$$

[16]Moreover, if public good quality varies across neighborhoods (e.g. due to peer effects in neighborhood schools), then further refinement of the distribution can be made.

A direct consequence of assumption 6 is that the housing demand functions above do not depend on α and g. Assumption 5 implies that we observe the joint distribution of income and housing for each community. If the function $V^b(y,p)$ satisfies standard integrability conditions, the function $V^b(y,p)$ is identified (up to a monotonic transformation) as long as we observe the joint distributions of housing and income for each community.[17]

Identification therefore focuses on $V^g(g)$. The main problem for identification 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 satisfy the following condition:

$$\alpha_{j+1}(y) \quad > \quad \alpha_j(y) \quad \forall j \tag{19}$$

is consistent with observed outcomes.

As a consequence of this property, we have the following result:

Proposition 3 *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 6;
(ii) $V(\alpha,y,g,p)$ implies the same housing demand functions as the true indirect utility function $V_0(\alpha,y,g,p)$;
(iii) $\alpha_{j+1}(y) > \alpha_j(y)$ $\forall j$,
there exists 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 functions $V_a(\alpha,y,g,p)$ 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.

[17]Technically speaking, we can only recover the demand functions at the observed housing prices which then imposes limits on identification of the indirect utility function.

Condition (iii) implies that

$$\alpha^a_{j+1}(y) \quad > \quad \alpha^a_j(y) \quad \forall j \tag{20}$$

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

$$F_{ia}(\alpha^a_j(y)|\, y) \quad \equiv \quad F_{i0}(\alpha^0_j(y)|\, y) \quad j = 1, ..., J \tag{21}$$

Then by construction, the observed equilibrium sorting of households by income within and among communities 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 distributions of income and housing are also the same for each community. Q.E.D.

We have seen that the locational equilibrium considered in this paper is only partially identified. Conditional on knowing the indirect utility function, we can identify J-1 points of the joint distribution of income and tastes. Point identification of the utility function $V_0(\cdot)$ and the distribution of households types $F_0(\cdot)$ is not feasible for general specifications of the model.

3.2 Semiparametric Identification

To obtain point identification, we need to impose stronger assumptions. We adopt a semiparametric framework and introduce a parametrization of the indirect utility function. We use a constant elasticity of substitution formulation to capture the trade-offs between the public good and the private goods components. For concreteness, we also adopt a form for $V^b(y, p)$ that implies constant price and income elasticities for housing.

Assumption 7 *The utility function is known up to a finite vector of parameters, θ, and* *takes the form:*

$$V(\alpha, y, g_j, p_j) \;=\; \left\{\alpha\, g_j^{\rho} \;+\; \left[e^{\frac{y^{1-\nu}-1}{1-\nu}}\; e^{-\frac{Bp_j^{\eta+1}-1}{1+\eta}}\right]^{\rho}\right\}^{\frac{1}{\rho}} \tag{22}$$

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

Assumption 7 implies that the set of households that are indifferent between adjacent communities is implicitly characterized by the following equation:

$$\alpha_j(y) \;=\; \left[e^{\frac{y^{1-\nu}-1}{1-\nu}}\right]^{\rho}\; \frac{Q(p_j) - Q(p_{j-1})}{g_{j-1}^{\rho} - g_j^{\rho}} \;\equiv\; \left[e^{\frac{y^{1-\nu}-1}{1-\nu}}\right]^{\rho}\; e^{K_j} \tag{23}$$

where $Q(p_j) = e^{-\,\rho\,\frac{Bp_j^{\eta+1}-1}{1+\eta}}$. Roy's identity implies that housing consumption is given by

$$h(p, y) \;=\; B\, p^{\eta}\, y^{\nu}. \tag{24}$$

Note that η is the price elasticity of the demand for housing, ν is the income elasticity and B is the scale parameter of the housing demand equation.

The discussion in the previous section directly implies the following proposition:

Proposition 4 *The following results hold for our semiparametric model:*

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

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:*

$$\left\{\rho \;\middle|\; \alpha_{j+1}(y|\rho) > \alpha_j(y|\rho) \quad \forall j\right\} \tag{25}$$

This set contains ρ_0.

Next we consider two cases that yield stronger identification results. The first case is based on the assumption that tastes and income are independently distributed in the population:

Assumption 8 *There exists at least one household type i for whom income and tastes are independently distributed, i.e. $f_i(\alpha, y) = f_i(\alpha)\ f_i(y)$.*

Under these additional assumptions, we can provide sufficient conditions for point identification of the model. The intuition is the following. Holding θ fixed, we can identify the marginal distribution of tastes by inverting the income distribution of an arbitrary community in our sample. For example, for community 1 we have

$$F_i(\alpha_1(y)) = \frac{f_{i1}(y)}{f_i(y)}\, n_{i1} \tag{26}$$

Equation (23) implies that $\alpha_1(y)$ is a smooth continuous function such that:

$$\begin{aligned} \lim_{y \to 0} \alpha_1(y) &= \infty \\ \lim_{y \to \infty} \alpha_1(y) &= 0 \end{aligned} \tag{27}$$

Thus there is a one to one mapping between the observed income distribution in community 1 and the distribution of the unobserved tastes. The model is identified if we can only predict the remaining income distributions at the correct value θ_0. The following proposition provides a necessary and sufficient condition for point identification under independence.

Proposition 5 *Suppose that (i) there exists one household type i for whom income and the taste for public goods are independently distributed; and (ii) for all $\rho < 0$, $\rho \neq \rho_0$, there exists a $j > 1$ such that $\frac{1}{\rho}[K_j(\rho) - K_{j-1}(\rho)] \neq \frac{1}{\rho_0}[K_j(\rho_0) - K_{j-1}(\rho_0)]$. Then the model is globally identified at (F_0, ρ_0).*

Proof: Fix α at an arbitrary level α_1. For any community $j > 1$ and any income level y, there exists a level of income $\tilde{y}$ such that

$$\begin{aligned} \ln(\alpha_1) &= K_j(\rho_0) + \rho_0 \frac{y^{1-v} - 1}{1 - v} \\ &= K_{j-1}(\rho_0) + \rho_0 \frac{\tilde{y}^{1-v} - 1}{1 - v} \end{aligned} \tag{28}$$

Independence of α and y then implies that for any arbitrary type i :

$$\begin{aligned} F_{i0}(\alpha_1) &= F_{i0}\left(e^{K_j(\rho_0) + \rho_0 \frac{y^{1-v}-1}{1-v}}\right) \\ &= F_{i0}\left(e^{K_{j-1}(\rho_0) + \rho_0 \frac{\tilde{y}^{1-v}-1}{1-v}}\right) \end{aligned} \tag{29}$$

Now consider an alternative structure (F_{ia}, ρ) not equal to (F_{i0}, ρ_0). For this alternative model to be observationally equivalent to (F_{i0}, ρ_0), we need that

$$\begin{aligned} F_{i0}\left(e^{K_j(\rho_0) + \rho_0 \frac{y^{1-v}-1}{1-v}}\right) &= F_{ia}\left(e^{K_j(\rho) + \rho \frac{y^{1-v}-1}{1-v}}\right) = \frac{\sum_{k=1}^{j} n_{ik} f_{ik}(y)}{f_i(y)} \\ F_{i0}\left(e^{K_{j-1}(\rho_0) + \rho_0 \frac{\tilde{y}^{1-v}-1}{1-v}}\right) &= F_{ia}\left(e^{K_{j-1}(\rho) + \rho \frac{\tilde{y}^{1-v}-1}{1-v}}\right) = \frac{\sum_{k=1}^{j-1} n_{ik} f_{ik}(\tilde{y})}{f_i(\tilde{y})} \end{aligned} \tag{30}$$

Equations (29) and (30) imply that there exists an α_2 such that:

$$\ln(\alpha_2) = K_{j-1}(\rho) + \rho \frac{\tilde{y}^{1-v} - 1}{1 - v} = K_j(\rho) + \rho \frac{y^{1-v} - 1}{1 - v} \tag{31}$$

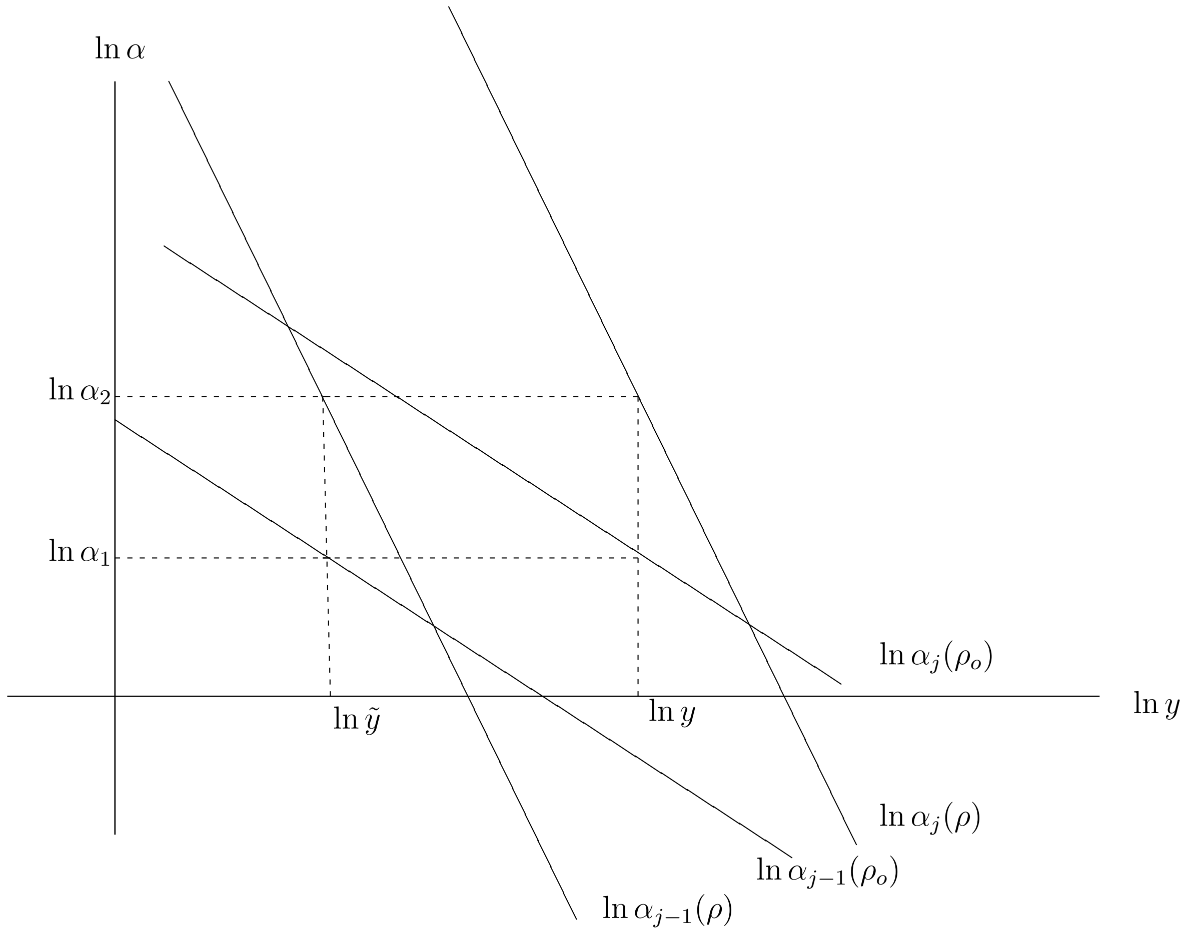
Equations (28) and (31) imply

$$K_j(\rho) - K_{j-1}(\rho) = \frac{\rho}{\rho_0}(K_j(\rho_0) - K_{j-1}(\rho_0)) \tag{32}$$

which contradicts condition (ii). Q.E.D.
Figure 1 provides a graphical illustration of the proof.

Condition (ii) is easily checked computationally. We computed an equilibrium with a large number of communities and found condition (ii) to hold. Indeed, we found that

Figure 1: Identification and Independence



$\frac{1}{\rho}[K_j(\rho) - K_{j-1}(\rho)]$ was monotone for one community in our example, which implies (ii).

Independence is admittedly a strong assumption. However, the basic idea behind the identification proof can also be used to establish identification of models that are more interesting.[18] Let us assume, for expositional simplicity, that $I = 2$. Let d_i be an indicator which is equal to one if $i = 2$ and zero otherwise. Let us also assume that the $F_1(\alpha|y)$ and $F_2(\alpha|y)$ satisfy the following shape restriction:

Assumption 9 *The distribution $F_i(\alpha|y)$ satisfies the following condition:*

$$\begin{aligned} F_i(\alpha|y)] &= Pr\{exp(d_i\beta(y) \;+\; \epsilon) \leq \alpha|\; x_i, y\} \\ &= Pr\{\epsilon \leq \ln(\alpha) \;-\; d_i\beta(y)|\; x_i, y\} \\ &= G(\ln(\alpha) \;-\; d_i\beta(y)|\; y) \end{aligned}$$

where $G(\cdot|y)$ denotes the conditional distribution of ϵ, and $\beta(y)$ is a function which characterizes the differences in the conditional means of the two distributions.

Substituting equation (33) into the equation characterizing household sorting yields:

$$G(\ln \alpha_j(y) \;-\; d_i\beta|y) \;=\; \frac{\sum_{k=1}^{j} n_{ik} f_{ik}(y)}{f_i(y)} \;\equiv\; G_{ij}(y) \tag{33}$$

Note that Proposition 1 implies that we can now identify $I*(J-1)$ points of the conditional distribution function of $G(\cdot|y)$ if we know ρ and β.

More importantly, we can construct bounds for the parameters $\beta_0(y)$ and ρ_0 and provide sufficient conditions for point identification. We then have the following result:

Proposition 6 *Suppose that there are three communities $j < l$ and k such that:*

$$G_{1j}(y) \;\leq\; G_{2k}(y) \;\leq\; G_{1l}(y) \tag{34}$$

[18]The independence assumption is also useful to establish identification of a model in which the household type is unobserved which is especially useful when modelling horizontal taste heterogeneity defined over a vector of local amenities and public goods.

The set of values $\beta(y)$ and ρ which are consistent with the inequalities above are given by:

$$\{(\beta(y), \rho\}| \; K_j(\rho) \leq K_k(\rho) \; - \; \beta \leq K_l(\rho)\} \tag{35}$$

Proof:

Fix income at some level y and consider the first inequality:

$$\begin{aligned} G(\ln \alpha_j(y)|y) \;&=\; G_{1j}(y) \\ &\leq\; G_{2k}(y) \\ &=\; G(\ln \alpha_k(y) \; - \; \beta(y)) \end{aligned} \tag{36}$$

Substituting equation (23) into the equation above yields:

$$G\left(\rho \, \frac{y^{1-\nu} - 1}{1-\nu} \; + \; K_j(\rho) \Big| \; y\right) \;\; \leq \;\; G\left(\rho \, \frac{y^{1-\nu} - 1}{1-\nu} \; + \; K_k(\rho) \; - \; \beta(y) \Big| \; y\right) \tag{37}$$

The monotonicity of $G(\cdot|y)$ then implies that

$$K_j(\rho) \;\; \leq \;\; K_k(\rho) \; - \; \beta(y) \tag{38}$$

A similar result applies for the second inequality. Q.E.D.

Proposition 6 allows the econometrician to construct bounds for the parameters β and ρ. In applications with a large number of communities, it is plausible to expect that condition (34) will not just hold for one set of three communities, but for a large number of community triples. In that case, we obtain a large number of inequality constraints that bound the main parameters of interest. We, therefore, expect that these bounds will be informative in application with many communities.

Proposition 6 also allows us to characterize conditions for point identification of $\beta(y)$ and ρ. If there exists at least two pairs of communities (j,k) and (m,n) such that

$$\begin{aligned} G_{1j}(y) &= G_{2k}(y) \\ G_{1m}(y) &= G_{2n}(y) \end{aligned} \tag{39}$$

Then we have

$$\begin{aligned} K_j(\rho) &= K_k(\rho) - \beta(y) \\ K_m(\rho) &= K_n(\rho) - \beta(y) \end{aligned} \tag{40}$$

which is a system of two nonlinear equations in two unknowns. If this system has a unique solution at ρ_0 and $\beta_0(y)$, then the parameters of the model are point-identified.[19] Note that point identification of the parameters of the utility function does not imply point identification of $G(\cdot|y)$. As before we can only identify values of the function that correspond to the points on the boundary loci of adjacent communities.

We have thus far provided in Propositions 5 and 6 two alternative approaches for identification of ρ. A third alternative is available, if we observe the locational equilibria of the same metropolitan area at two successive points of time. Suppose that preferences remain constant, but the income distribution of the metropolitan area changes between the two time periods. It is then straightforward to derive sufficient conditions for identification of the model. The intuition is the following. Conditional on ρ, the distributions $F_i(\alpha, y)$ are identified using the equilibrium allocation observed in the first period. We can then predict the equilibrium in the second period as a function of ρ. Since the distribution of income changes between the two periods, the equilibrium in period 2 will be different from the equilibrium in period 1. Thus the observed sorting by income will be distinctly different in both periods. We can then derive results similar to one in Proposition 6.

[19] As with Proposition 6, uniqueness is not guaranteed, in general, but it is readily checked computationally.

4 Estimation

4.1 First Stage

The proofs of identification of the model considered in this paper are constructive and can be used to devise new estimators for the underlying distribution of household types.[20] To derive these new estimators, it is desirable to relax assumption 5. Public good provision may not be perfectly observed by the econometrician. In most applications, it is likely that we will observe some measures of public good provision. However, our observed measures may be subject to measurement error.[21] Suppose, we we do not observe g_j, but we observe, $\tilde{g}_j$, which is given by:

$$\tilde{g}_j \;=\; g_j + \epsilon_j \tag{41}$$

where ϵ_j denotes measurement error. The ascending bundles property in Proposition 1 implies that the levels of public good provision are monotonically increasing in the (price) ranks of the communities. Let us denote the rank of community j by r_j. Hence ascending bundles implies that in equilibrium the following equation holds

$$g_j \;=\; g(r_j) \tag{42}$$

for some unknown monotonically increasing function $g(\cdot)$. Substituting equation (42) into equation (41), we obtain

$$\tilde{g}_j \;=\; g(r_j) + \epsilon_j \tag{43}$$

Furthermore suppose that $E[\epsilon_j|r_j] = 0$, i.e. the error term in equation (41) is conditionally independent of the rank of a community. In that case $g(r_j)$ is nonparametrically identified.

[20] We assume that we have consistent estimators of B, η and ν. Estimation of these parameters is straightforward and can be done prior to estimating the other parameters and functions of the model.

[21] One of the main advantages of parametric estimators such as those proposed by Berry, Levinsohn, and Pakes (1995) or Epple and Sieg (1999) is that they allow for unobserved heterogeneity among communities.

In the application considered in section 5, we estimate the function $g(\cdot)$ using locally-linear kernel regressions as suggested by Fan (1992).[22] Given a smoothing parameter h and a kernel function $K(\cdot)$, the estimator has an asymptotic bias of $\frac{1}{2}h^2 g''(r) \int u^2 K(u) du$ and variance of $\frac{1}{hN} \frac{\sigma^2(r)}{f(r)} \int K^2(u) du$, which can be estimated consistently to compute plug-in confidence bands.[23]

For the locational equilibrium model to be well defined, we also need that the community specific intercepts are monotonically increasing: $K_1 < ... < K_J$. A necessary but not sufficient condition for that 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 public good provisions are sufficiently large relative to the differences in observed housing prices, then the intercepts are also monotonically increasing functions. It is therefore desirable to impose these curvature restrictions in estimation. Moreover, by testing whether these curvature restrictions hold in the data, we can devise a specification test for our model.

One way to impose these curvature restriction is to use isotone-kernel regression estimators proposed by Mammen (1991). These estimators use a two step procedure to estimate a monotonic function. First, a nonparametric estimator $\hat{g}(\cdot)$ is obtained, using for example the local linear kernel estimator discussed above. This function may not be monotonically increasing. In the second step, the estimated function is projected onto a space of monotonically increasing functions. The estimator is defined as:

$$\tilde{g}(r) = \text{argmin}_{g \in G} \int (g(r) - \hat{g}(r))^2 \, dr \tag{44}$$

where G denotes the class of shape restricted functions. We can use a similar approach to impose the curvature restrictions implied by our model. To see how this works, define $m = -g^\rho$ and $\hat{m} = \hat{g}^\rho$. The curvature restrictions then imply that $m(r)$ is sufficiently

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

[23]The estimation procedure can also be extended to account for multiple public goods as long as public good provision satisfies an index assumption. Suppose we observe a vector of community specific amenities x_j. Let us assume that household preferences only depend on the linear index $g_j = x_j'\gamma + \epsilon$. In that case we need to combining the techniques discussed in this section with those suggested by Robinson (1988). We need to normalize the coefficient of one the components in the index to achieve identification.

monotonically increasing in r. Our constrained estimator of the function is then obtained by minimizing the following objective function:

$$\min \sum_{j=1}^{J} (m_j - \hat{m}_j)^2 \tag{45}$$

subject to the constraints that:

$$\begin{aligned} m_{j-1} &< m_j - \delta_2 \\ \frac{1}{Q_{j-1} - Q_{j-2}}(m_{j-1} - m_{j-2}) &\leq \frac{1}{Q_j - Q_{j-1}}(m_j - m_j) - \delta_1 \end{aligned} \tag{46}$$

for some 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. Standard errors and confidence bands are computed using bootstrap techniques.[24]

The constrained estimator of the function $g(\cdot)$ depends on ρ because the constraints are functions of ρ. In principle, one can implement this estimator for any choice of ρ. In our application, we have found that the estimated constrained functions are very similar for all reasonable values of ρ. If the estimated function $g(\cdot)$ satisfies the the constraints for one value of ρ, it also satisfies these constraints for almost all other plausible values of ρ.

4.2 Second Stage

The constrained estimator of $g(\cdot)$ directly implies estimators of the boundary indifference loci $\alpha_j(y|\rho)$ which are well behaved. We denote these estimators by $\hat{\alpha}_j(y|\rho)$. Suppose we observe the empirical market shares and income distributions for each community, $\{n_{ij}^N, f_{ij}^N(y)\}_{j=1}^J$, where N denotes the relevant sample size. We can then proceed and estimate the conditional distribution of tastes. Following the discussion of identification in the previous section, a nonparametric estimator of the $J-1$ points of the conditional

[24] Bootstrap techniques are discussed in Efron and Tibshirani (1993) and Hall (1994).

distribution of tastes given income, $\hat{F}_i^N(\alpha|\ y)$, is then given by:

$$\hat{F}_i^N(\hat{\alpha}_j(y|\rho)|\ y) \quad = \quad \sum_{k=1}^{j} \frac{f_{ik}^N(y)}{f_i^N(y)}\, n_{ik}^N \quad j = 1, ..., J-1 \tag{47}$$

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

$$\sqrt{hN}\ \left(\sum_{k=1}^{j} \frac{f_{ik}^N(y)\ n_{ik}^N}{f_i^N(y)} \ - \ \sum_{k=1}^{j} \frac{f_{ik}(y)\ n_{ik}}{f_i(y)}\right) \xrightarrow{d} N(0, \sigma_j^2(y)) \tag{48}$$

where the asymptotic variance $\sigma_j^2(y)$ can be computed easily from the variances of the density estimators using the delta-method.

Next consider the semiparametric model with two observed types discussed at the end of the previous section. Notice that if the model is correctly specified, we have:

$$F_1(\alpha_j(y|\ \rho)|y) \quad = \quad F_2(\alpha_j(y|\rho)\ +\ \beta(y)\ |y) \tag{49}$$

For a given grid $y_1, ..., y_B$, define an estimator for $\beta = (\beta(y_1), ..., \beta(y_B))$ and ρ as follows

$$(\hat{\beta}^N, \hat{\rho}^N) \quad = \quad \operatorname{argmin} \sum_{b=1}^{B} \sum_{j=1}^{J-1} \Big(F_1^N(\alpha_j(y_b|\rho)|y_b) \ - \ F_2^N(\alpha_j(y_b|\rho)\ +\ \beta(y_b)|y_b)\Big)^2 \tag{50}$$

For any finite J this estimator is technically not a feasible estimator since $F_2^N(\alpha_j(y_b|\rho) + \beta(y_b)|y_b)$ may not be identified. However in an application with a large number communities there will exist a community k such that

$$F_2^N(\alpha_k(y_b|\rho)\ |y_b) \leq F_2^N(\alpha_j(y_b|\rho) +\ \beta(y_b)|y_b) \leq F_2^N(\alpha_{k+1}(y_b|\rho)\ |y_b) \tag{51}$$

As J grows large, the difference $F_2^N(\alpha_{k+1}(y_b|\rho)\ |y_b) \ - \ F_2^N(\alpha_k(y_b|\rho)\ |y_b)$ gets small for a large number of community pairs. Thus in applications with a large number of communities the error which arises when approximating the value of $F_2^N(\alpha_j(y_b|\rho) +\ \beta(y_b)|y_b)$ is negligible for many j.

The intuition behind the estimator in equation (50) is the following. Given a value of ρ the distributions $F_1(\alpha|y)$ and $F_2(\alpha|y)$ are identified (at $J-1$ points). Assumption 9 implies that $\beta(y)$ is a vertical shift which characterizes the differences in means of the two conditional distributions. Thus conditional on ρ, $\beta(y)$ can be estimated as the average difference between the conditional distribution functions of both household types. Changing ρ implies a nonlinear transformation of the support of both distributions since $\alpha_j(y|\rho)$ is a nonlinear function of ρ.

In summary we have developed a new semiparametric estimator for the model with observed household types. In the first step of the estimation procedure, we nonparametrically estimate a function of public good provision $g(r)$. We have shown how to impose monotonicity and curvature restrictions on this function that are implied by the underlying economic theory. In the second stage, we estimate the underlying distribution of income and tastes by inverting the observed income distributions of the communities in the sample.

5 Data

Our application focuses on Allegheny County in Western Pennsylvania, 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.

We have obtained a detailed data set on the 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):

$$r_{jn} = [(1 - \tau_y)(i + \tau_p) + \beta + m + \delta - \pi]\, v_{jn} \tag{52}$$

with $\tau_y = 0.15$, $i = 0.079$, $\beta = 0.04$, $m - \delta = 0.02$, and $\pi = .02866$.

To construct our 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.

We construct an education index based on the PSSA, a test administered in all public schools in Pennsylvania in the school year 1999-2000. The PSSA consists of tests in math and reading administered in grades 5, 8, and 11. 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 participations 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.

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. Summary statistics are reported in Table 1.

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
Violent Crime Index	392	575	0	4771
Property Crime Index	2358	2809	0	27368
Total Crime Index	690	888	0	8197
Rush hour travel time	24.22	11.15	1	57
Property Tax Rate	0.0202	0.0026	0.0140	0.0283
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

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:

$$\tau_j = \tau_j^p + \tau_j^y \frac{\bar{y}_j}{\bar{v}_j} \tag{53}$$

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

$$\ln r_{jn} = \sum_{j=1}^{J} I_j \ln p_j + \delta z_n + \varepsilon_n \tag{54}$$

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.

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

$$\ln r_{jq} - \ln p_j = \ln \beta + \nu \ln y_q + \eta \ln p_j + \varepsilon_{jq}^{h} \tag{55}$$

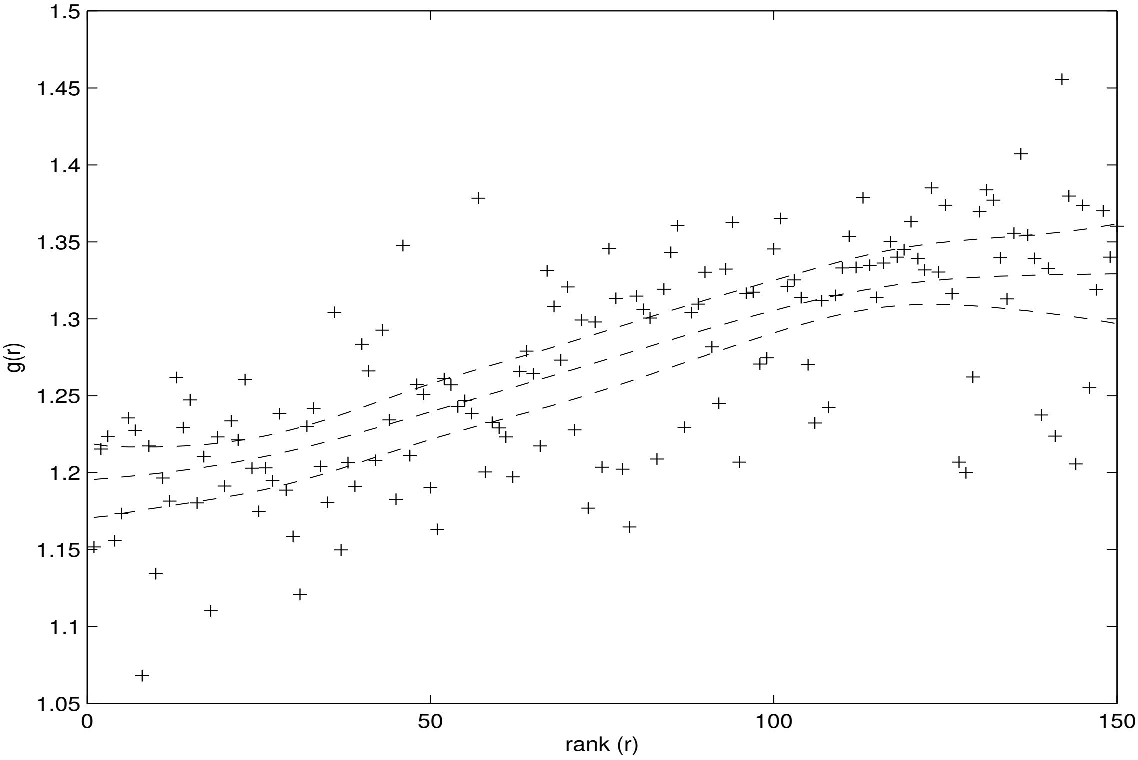
Here, r_{jq} represents the housing quantiles and y_{jq} represents the income quantiles. We use the 10% though 90% deciles in our estimation. The estimation results for the housing demand model yield an estimate for ν of 0.784 (0.017). The price elasticity η is estimated at -0.514 (0.027) and the intercept B is equal to 1.161 (0.182). The adjusted R^2 of the regression is 0.877.

6 Estimation Results

We estimate the function $g(r)$ using locally linear kernel regressions using cross validation to select the bandwidth parameter. 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 $K_j \leq K_{j+1}$ for any reasonable value of ρ.[25]

[25] 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).

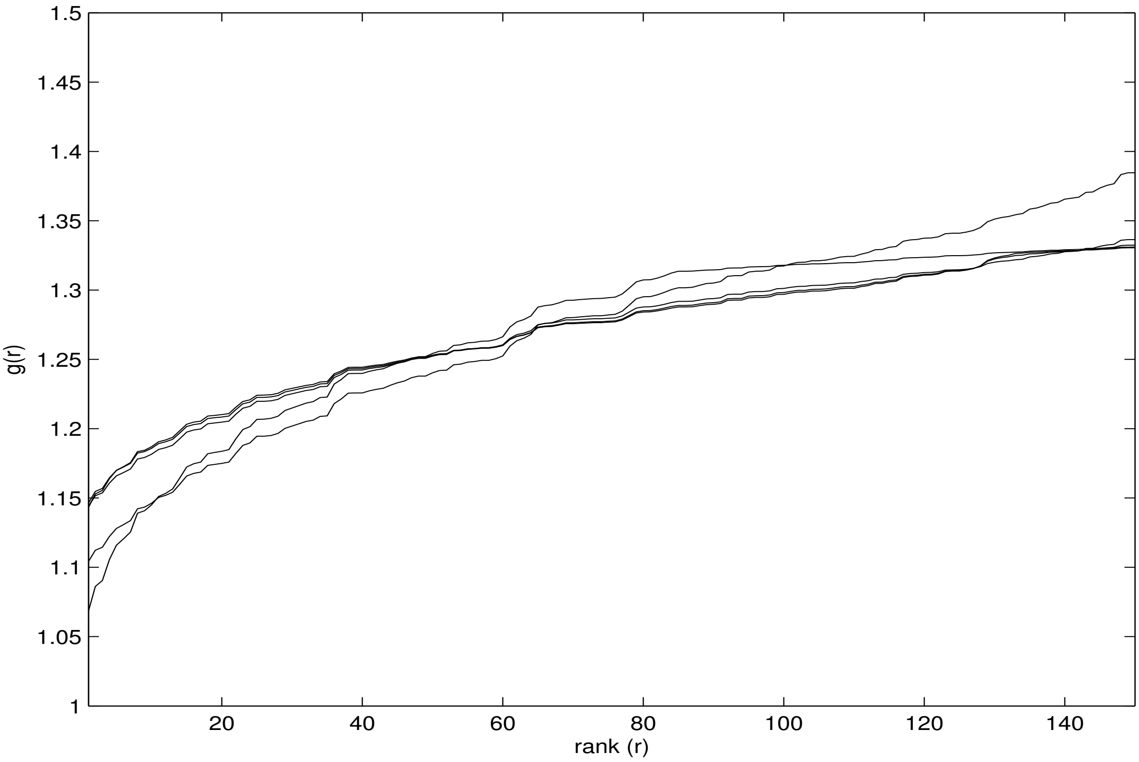
Figure 2: The First Stage: Estimation of $g(r)$



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

We, therefore, impose the curvature restriction implied by the ascending bundles 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 -1.0.[26] The results of these different estimators are plotted in Figure 3. For values of ρ ranging from -.2 to -.8, we find that the functions are similar. The restricted function falls between the 95% confidence bounds of the unrestricted function for most of its range. Using smaller values than -.8 or larger values than -.2 primarily affects the predicted values for g for the lowest and highest priced communities.

Figure 3: The First Stage: Constrained Estimation of $g(r)$



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

All of the functions plotted in Figure 3 yield estimates of the community specific inter-

[26]These values are the most plausible candidates for ρ based on previous (parametric) studies. For example, when we estimate the baseline model using the parametric approach suggest in Epple and Sieg(1999) using this data set, the estimate for ρ is equal to -.198.

cepts, K_j, that ascend not only for the value of ρ used in estimation, but also for most other plausible values of ρ. Based on these results, we therefore construct a constrained function of $g(\cdot)$ which satisfies the relevant inequality constraints for all plausible values of ρ. We use this function in the second stage of the estimation procedure.

The second stage of the estimation procedure is based on the observed 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 empirical probabilities which also play a large role in the formal estimation:

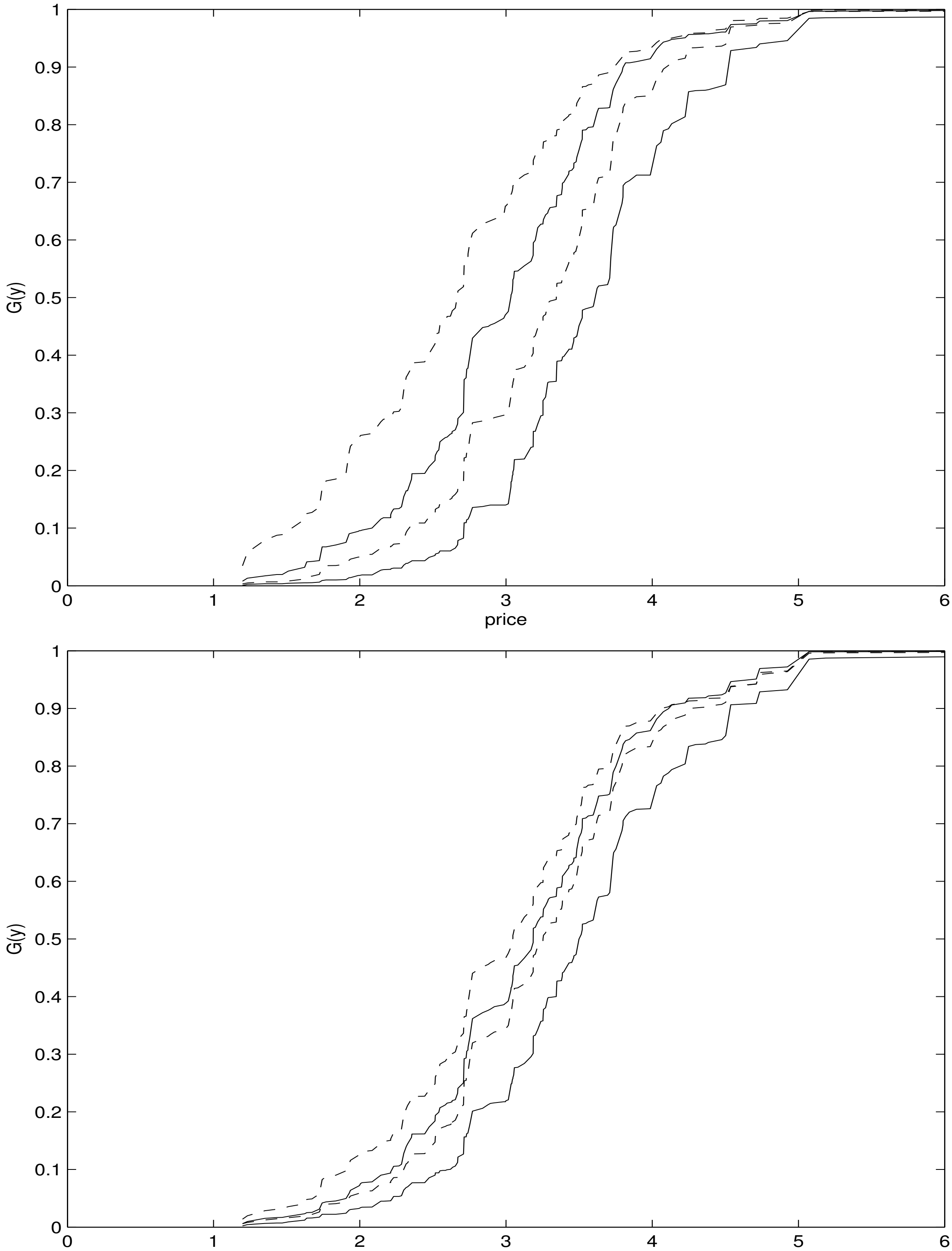
$$\sum_{k=1}^{j} \frac{f_{ik}^N(y)}{f_i^N(y)} \, n_{ik}^N$$

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. We estimate the densities for income in the each community using histogram estimators which are based on the 16 bins that the U.S. Census uses to report empirical income distributions.

We plot these probabilities against the estimated housing prices in Figure 4. The upper (lower) panel reports the plot for households with children (without children). We include plots 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.

Figure 4 provides some interesting new insights. Households with children seems to be more responsive to differences in housing prices (and local public good provision) than households without children. Consider low income households with annual income of \$19,330. 26 % of households with children live in the 20 cheapest communities. Only 13 % of households without children live in these communities. The same pattern holds for richer households

Figure 4: Sorting of Households by Income



The top panel shows households with children. The bottom panel shows households without children.

with annual income of \$102,239. 59 % of households with children live in the 50 most expensive communities. Here the corresponding number for households without children is 53 %. We thus conclude that the sorting of households with children exhibits more stratification by income than the observed sorting of households without children. The vertical difference between the curves are significantly larger for households with children than households without children. We also find that households with children and income levels below the mean metropolitan income are more likely to live in cheaper communities than households without children. The opposite is true for households with high levels of incomes. High income households with children have stronger preferences for high price (and high amenity) communities than households without children.

We then implement the second stage of the estimation procedure. We evaluate the conditional distribution of tastes at the four income levels discussed above. Table 2 reports the parameter estimates for ρ and $\beta(y)$ for the sample of 150 communities. We also consider a smaller subsample that consists of the first 120 communities. In this subsample, we exclude the 30 most expensive communities in our sample. The results for this smaller subsample provide a robustness check. We want to make sure that the parameter estimates are not driven by the upper tails of the conditional distribution functions. Table 2 also reports standard errors which are computed using a bootstrap algorithm which samples from the underlying community specific income distributions. We used 50 repetitions to computed the bootstrap errors. The objective function is minimized using a grid search algorithm.

Consider the results for the full sample. We find that the parameter estimate of ρ is -.58 with an estimated standard deviation of .16. The coefficient estimates of $\beta(y)$ range from -.21 for low income households to .05 for high income households. This finding indicates that low income households with children have significantly lower tastes for public goods than households without children. In contrast high income households with children have, on average, stronger tastes for local public goods than households without children. Recall that these parameter estimates are identified of the observed differences in the distribution of household types by income among the set of communities in the sample. Our estimates reflect the fact that lower income households with children are more likely to live in cheaper

Table 2: Estimation Results

parameters	full sample	subsample
ρ	-0.580	-0.590
	(0.160)	(0.112)
$\beta(19330)$	-0.210	-0.217
	(0.058)	(0.042)
$\beta(37921)$	-0.089	-0.082
	(0.021)	(0.017)
$\beta(66247)$	0.021	0.030
	(0.008)	(0.007)
$\beta(102239)$	0.050	0.068
	(0.008)	(0.013)
objective function	0.352	0.225

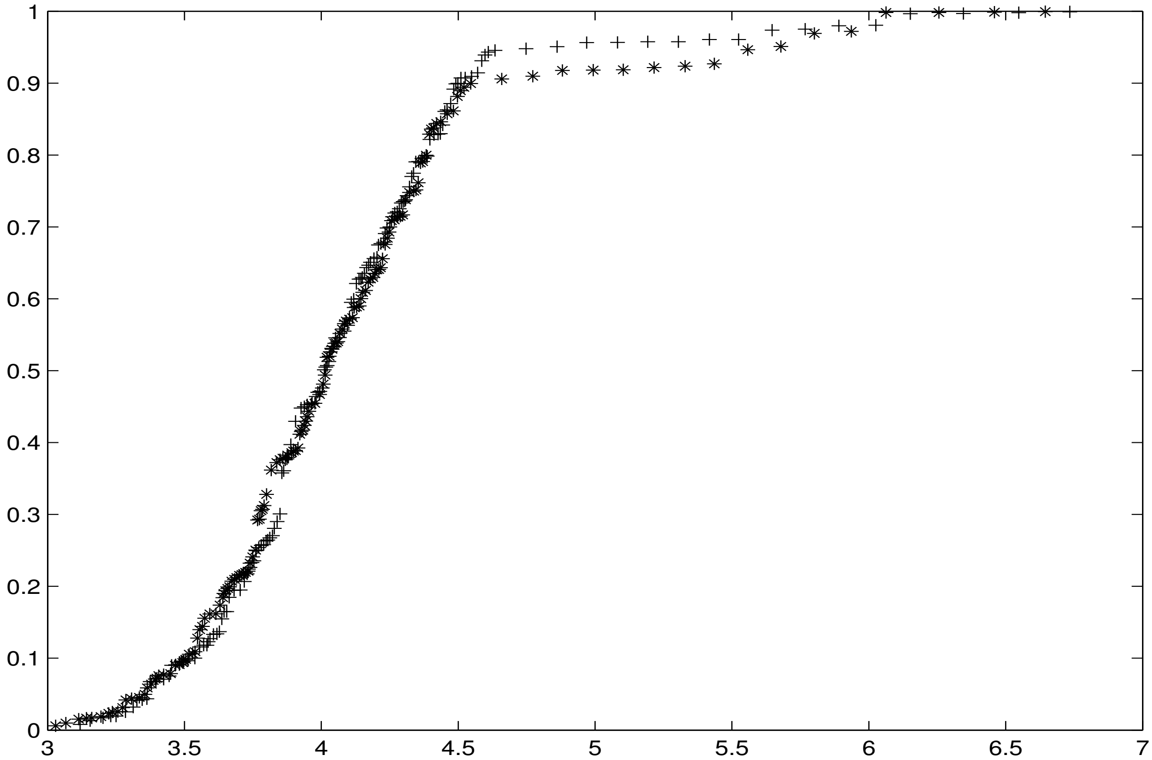
Estimated standard errors are given in parentheses.

communities with low levels of public good provisions than households without children. The opposite is true for high income households. Comparing columns I and II shows the results are not driven by the tail behavior of the distribution functions.

Next we consider the fit of the model. Figure 5 plots the conditional distributions of tastes for both household types evaluated at the median level. After controlling for the shift in means. A $'+'$ indicates households without children and a $'*'$ indicates household with children. Under the null hypothesis, the two conditional should line up perfectly after controlling for the shift in means. Figure 5 shows that the fit of the model is good. For values of the distribution function below 0.9, the two distribution functions are very close. The main difference between the two distribution functions occurs in the upper tail. Note that this upper tail of the distribution is harder to identify since there are only a few communities in our sample which identify this part of the distributions. Moreover, the housing price estimates for the communities may be biased upward because we may not measure housing consumption well for these high income communities. We have also performed similar plots for lower and higher income levels. We find that the results are qualitatively the same.

Finally, we note that the estimator provides, by construction, a perfect fit of the ob-

Figure 5: Second Stage



The figure plots the two type specific conditional distribution function of tastes for households with median income.

served sorting of households without children by income within and across communities. These households account for approximately 73.3% of all households in the Pittsburgh metropolitan area. Given the good fit of the second stage estimator, we also closely predict the observed sorting patterns of households with children. The semi-parametric estimator thus significantly improves the fit of the model compared to previously used parametric estimators.

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. Given knowledge of the indirect utility function, we can identify $J-1$ points of the distribution of tastes conditional on income for each household type. This result shows that there is a linear relationship between the number of communities in the choice set and the number of points of the conditional distribution of tastes that are nonparametrically identified. Joint identification of the distribution of household types and the indirect utility function is more problematic. We have shown that plausible separability assumptions allow us to identify the sub-utility function which models tastes over private goods. Moreover, we can construct informative bounds for the sub-utility function which models tastes for public goods. To obtain stronger results, we adopt a semiparametric approach using a parametric specification of the indirect utility function. We have considered two alternative scenarios and have provided sufficient conditions for point identification in each scenario. The first case is based on the assumption that tastes and income are independently distributed for at least one household type. The second case imposes additional restrictions on the shape of the conditional taste distributions of the two observed household types.

The proofs of 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 algorithms. 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 the properties of this estimator using an application that focuses on sorting of households with and without children across municipalities in Allegheny County. We have provided summary statistics which document 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 differentiated product models including 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

Athey, S. and Haile, P. (2002). Identification in Standard Auction Models. *Econometrica*, *70 (6)*, 2107–2140.

Bajari, P. and Benkard, L. (2002). Demand Estimation With Heterogenous Consumers and Unobserved Product Characteristics: A Hedonic Approach. Working Paper.

Bayer, P., McMillan, R., and Reuben, K. (2003). 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. (2004). Myopic Voting and Local Public Good Provision. Working Paper.

Chesher, A. (2003). Identification in Nonseparable Models. *Econometrica*, *71 (5)*, 1405–1441.

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. (1987). Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products. *Journal of Political Economy*, *95*, 59–80.

Epple, D., Filimon, R., and Romer, T. (1984). Equilibrium Among Local Jurisdictions: Towards an Integrated Approach of Voting and Residential Choice. *Journal of Public Economics*, *24*, 281–304.

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. (1996). Income Distribution, Communities, and the Quality of Public Education. *Quarterly Journal of Economics*, *111 (1)*, 135–164.

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.

Ferreyra, M. (2003). Estimating the Effects of Private School Vouchers in Multi-District Economies. Working Paper.

Gallant, R. and Nychka, D. (1987). Semi-Nonparametric Maximum Likelihood Estimation. *Econometrica*, *55(2)*, 363–390.

Goodspeed, T. (1989). A Reexamination of the Use of Ability-to-Pay Taxes by Local Governments. *Journal of Public Economics*, *38 (3)*, 319–342.

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

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.

Hotz, J. and Miller, R. (1993). Conditional Choice Probabilities and Estimation of Dynamic Models. *Review of Economic Studies*, *60*, 497–529.

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

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.

Manski, C. and Tamer, E. (2003). Inference on Regressions with Interval Data on a Regressor or Outcome. *Econometrica*, *71*, 519–546.

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

Nechyba, T. (1997a). Existence of Equilibrium and Stratification in Local and Hierarchical Tiebout Economies with Property Taxes and Voting. *Economic Theory*, *10(2)*, 277–304.

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

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.

Sieg, H., Smith, V. K., Banzhaf, S., and Walsh, R. (2002). Interjurisdictional Housing Prices in Locational Equilibrium. *Journal of Urban Economics*, *50*, 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.

Tiebout, C. (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy*, *64 (5)*, 416–424.

Walsh, R. (2002). Analyzing Open Space Policies in a Computable Locational Equilibrium Model. Dissertation, Duke University.