Partial Identification of Preferences in a Dual-Market Sorting Equilibrium

Nicolai V. Kuminoff *†

ABSTRACT: This paper develops a new structural estimator that uses the properties of market equilibrium, together with information on households and their observed location choices, to partially identify horizontally differentiated preferences for a vector of local public goods. The estimation is consistent with equilibrium capitalization of local public goods and recognizes that job and house location choices are interrelated. By using set identification to distinguish the identifying power of restrictions on the indirect utility function from the identifying power of assumptions on the distribution of preferences, the estimator provides a new perspective on models of sorting equilibria. The estimator is used to recover distributions of the marginal willingness-to-pay for improved air quality in Northern California's two largest population centers: the San Francisco and Sacramento metropolitan areas. The average marginal willingness-to-pay increases by up to 110% when job opportunities are included as a dimension of location choice.

June 2012

KEYWORDS: Sorting, locational equilibrium, partition, dual-market.

* Assistant Professor, Department of Economics, Arizona State University, kuminoff@asu.edu

† I thank Pat Bajari, Steven Berry, Ken Chay, Paul Fackler, Roger von Haefen, Mike Keane, Ray Palmquist, Chris Parmeter, Dan Phaneuf, Jaren Pope, Chris Timmins, three anonymous referees, and especially V. Kerry Smith for helpful comments and suggestions on this research. I also thank seminar participants at Arizona State University, the EPA National Center for Environmental Economics, North Carolina State University, University of California, Berkeley, University of California, Davis, University of Minho, University of Rochester, University of Wyoming, University of Tennessee, and Virginia Tech. Support from the National Science Foundation is gratefully acknowledged.

"There is no way in which the consumer can avoid revealing his preferences in a spatial economy. Spatial mobility provides the local public-goods counterpart to the private market's shopping trip." --Charles Tiebout (1956)

1. INTRODUCTION

Nearly 60 years ago, Charles Tiebout suggested that consumers reveal their preferences for local public goods by the residential locations they choose. Epple and Sieg (1999) were the first to implement Tiebout's logic by using the properties of equilibrium in the housing market to estimate households' heterogeneous preferences for a composite local public good. In their analysis, households choose where to live based on their (exogenous) income and their preferences for the unique bundle of local public goods provided by each of a discrete set of urban communities. Households are depicted as differing in their tastes for the bundle of public goods, but they are restricted to evaluate its constituent elements in the same way. This feature, labeled vertical differentiation, implies all households agree on a single ranking of communities by an index of the public goods they provide.

Relaxing vertical differentiation is important because it is reasonable to expect that different households will evaluate components of a vector of local public goods quite differently. For example, households with school age children may be more concerned about school quality while retirees may place more emphasis on climate and other environmental amenities. While several microeconometric strategies have been proposed for the situation where households differ in their relative preferences (i.e. horizontal differentiation), none have used the properties of a market equilibrium to recover preferences in a way that is consistent with equilibrium capitalization of local public goods (Starrett [1981] and Scotchmer [1985]).

Equally important is the need to recognize that working households make two related location choices—the choice of a house and the choice of a job. The few existing studies that model adjustment in both markets use reduced form models that restrict preferences to be homogeneous and limit the analysis to marginal changes (e.g. Roback

[1982]) and Blomquist et al. [1988]).[1]

This paper describes a new structural estimator that meets both objectives, while nesting Epple and Sieg's (1999) model as a special case and extending Epple, Peress, and Sieg's (2005) semiparametric identification strategy to the case where households differ in their relative preferences for multiple public goods. More precisely, the new estimator is based on the information provided by location choices in a market equilibrium derived from households that have horizontally differentiated preferences for public goods and differ in their job skill. It recognizes that observed location choices provide set identification of the heterogeneous preference parameters. That is, the estimator recovers a set of values for the parameters that describe how local public goods contribute to sorting behavior. To attach values from this set to the population of households requires additional assumptions about the distribution of each preference parameter. A key feature of the new estimator is that it uses the set identification logic to distinguish the identifying power of structural restrictions on the indirect utility function from the identifying power of maintained assumptions about the distribution of preferences.

To evaluate the implications of introducing a joint job-house choice and heterogeneous relative preferences into an equilibrium sorting model, the new "dual-market" estimator and Epple and Sieg's model are both used to recover preferences for public goods in Northern California's two largest population centers: the San Francisco and Sacramento Consolidated Metropolitan Statistical Areas. This region is divided into 122 housing communities and 8 work destinations, and each (community, worksite) pair is assigned a price of housing, a set of public goods, a set of wage rates, and a commute time. Both models are used to explain the location choices made by households in each of 22 occupational categories, where wage options differ for each category in the dual-market case. Results from the estimation are used to construct distributions of the marginal willingness-to-pay for improved air quality. Moving from Epple and Sieg's model to the new "dual-market" framework increases estimates for the average per/household

[1] Rosen (1979) suggested that because households can make adjustments in both markets, we should expect both wage rates and house prices to reflect the demand for local public goods. Despite empirical evidence in support of Rosen's insight, most economists have focused exclusively on the housing component of location choice as a means to infer households' valuation of amenities.

marginal willingness-to-pay by as much as 110%.

Section 2 reviews the logic of Tiebout sorting in the context of Epple and Sieg's (1999) framework and discusses how structural restrictions allow preferences for public goods to be inferred from observed house locations. Then the choice set is expanded to include the labor market and a single-crossing restriction is used to characterize equilibrium sorting behavior. Section 3 describes the empirical model and the estimator. Then section 4 introduces the data and section 5 compares the results from implementing the new estimator to the results from two special cases—the Epple-Sieg model and an intermediate version of the model that admits horizontal differentiation but treats wage income as exogenous. After interpreting the results, section 6 concludes.

2. THEORY

Tiebout's locational sorting model assumes, ceteris paribus, heterogeneous households select a community based on its local public goods. Suppose the urban landscape can be divided into a finite set of J housing communities, each of which differs in its price of housing (p_j) and in its exogenous provision of local public goods such as school quality, crime, and environmental amenities. Households differ in the relative importance they assign to each public good. Let γ represent relative preferences for the different components of a composite index of public good provision, and let $\bar{g}_j(\gamma)$ represent composite provision of public goods in community j as perceived by a $\gamma-$type household. Each household chooses the community that maximizes its utility, given its exogenous income (y) and its preferences (α) for the composite public good relative to private goods. For heuristic purposes, utility maximization can be depicted as a two-stage problem, where each household first determines the optimal quantities of housing and numeraire in every community and then chooses the community that maximizes its utility. The first stage is shown as equation (1).

(1) $$\max_{(h,b)} U[\bar{g}(\gamma), h, b, \alpha] \quad subject\ to \quad ph = y - b\,.$$

Conditional on a community, households choose quantities of housing (*h*) and a composite private good (*b*) to maximize their utility subject to the budget constraint. Assume that zoning does not constrain housing construction. Then households can purchase any quantity of housing at the market price in each community, in which case preferences can be restated using the indirect utility function in (2).

$$(2) \qquad V[\bar{g}(\gamma), p, \alpha, y] = U[\bar{g}(\gamma),\ h(\bar{g}(\gamma), p, \alpha, y),\ \ y - ph(\bar{g}(\gamma), p, \alpha, y),\ \ \alpha].$$

Assuming households are price-takers and can move freely between communities, each household will choose the community that maximizes its well-being, given income and prices.

2.1. *Identifying Heterogeneous Preferences from Structural Restrictions*

Estimable models of Tiebout sorting rely on two types of structural restrictions to point-identify households' preferences based on their observed location choices. First, a parametric indirect utility function is selected. Second, a parametric distribution is specified for each preference parameter in that function used to characterize household heterogeneity. Each restriction makes a different type of contribution to the identification.[2]

Distributional assumptions compensate for discreteness in the choice set. When household *i* chooses *j* from a finite set of communities, utility maximization is characterized by the set of inequalities in equation (3).

$$(3) \qquad V[\bar{g}_{,ij}(\gamma_i), p_j, \alpha_i, y_i] \geq V[\bar{g}_{i,k}(\gamma_i), p_k, \alpha_i, y_i]\ ,\ \ \forall\ k = 1,\ldots, J\ .$$

Given a parametric form for the indirect utility function, the inequalities provide set identification of the heterogeneous preference parameters. It must be the case that $(\alpha_i, \gamma_i) \in A_{i,j}$, where $A_{i,j} = \{\ (\alpha_i, \gamma_i) : (\alpha_i, \gamma_i)\ \textit{satisfies}\ (3)\ \}$. In words, the choice of

[2] While discrete-choice studies typically rely on parametric assumptions for identification, it is also possible to identify the distribution of preferences nonparametrically if one is willing to impose a set of independence assumptions on the joint distribution of observed and unobserved variables together with a set of monotonicity and separability restrictions on the utility function (Briesch et al. [2007]).

community j reveals only that household i's preferences lie somewhere in the $A_{i,j}$ set. Imposing a distribution on (α, γ) allows the analyst to identify the density of preferences within $A_{i,j}$.

To illustrate the role of each type of restriction in identifying preferences, consider a specific example using the following CES indirect utility function:

$$V[g(\gamma), p, \alpha, y] = \left\{ \alpha_i (\bar{g}_{i,j})^{-.01} + .79[\exp(4y_i^{.25} - 54p_j^{.04})]^{-.01} \right\}^{\frac{1}{-.01}},$$

$$\text{where } \bar{g}_{ij} = \gamma_{i,air} AIR_j + \gamma_{i,school} SCHOOL_j.[3]$$

The first term represents utility from public goods, and the second term represents utility from the private good component of housing. Households differ in their income and in their preferences for a linear index of two public goods that differentiate communities, air quality and school quality. There are two components of preference heterogeneity. Households differ in the relative weights they assign to each public good in the index ($\gamma_{i,air}, \gamma_{i,school}$) and in the overall strength of their preferences for public goods relative to private goods (α_i). The weights are assumed to sum to 1 so that α_i represents a scaling parameter on the strength of preferences. Suppose households maximize their utility by sorting among the following four communities: $\{AIR_1 = 1.25, SCHOOL_1 = 1.25, p_1 = 1.00\}$, $\{AIR_2 = 1.85, SCHOOL_2 = 1.65, p_2 = 1.25\}$, $\{AIR_3 = 1.66, SCHOOL_3 = 1.86, p_3 = 1.26\}$, and $\{AIR_4 = 2.00, SCHOOL_4 = 2.00, p_4 = 1.50\}$, where higher values for *AIR* and *SCHOOL* indicate higher quality.

To see how the form of the indirect utility function provides set identification of preferences, first consider Epple and Sieg's (1999) vertically differentiated model. In this case all the variation in tastes can be condensed into a single heterogeneous parameter that ranks locations by "quality". The CES utility function simplifies to this case when

[3] This CES function provides the basis for the subsequent structural model. Specifically, it is the indirect utility function from equation (12) with $\beta = 2$, $\eta = -.963$, $\nu = .75$, $\phi = 0$, $w = 0$, and $\rho = -.01$. If the weights in the public goods index are constant, it reduces to the form of the indirect utility function in Epple-Sieg (1999).

households are constrained to have the same relative preferences for the two public goods. For example, let the weights be: $(\gamma_{i,air}, \gamma_{i,school}) \equiv (0.48,\ 0.52)$, $\forall i$. With constant weights, all households agree on a common ranking of communities by the public goods index, and sort according to their income and α_i . By conditioning on income, the system in (3) can be solved for the bounds of the α_i sets that rationalize each location choice. At *y=$50,000*, the partition of α corresponds to:

α	$A_{i,1}$		$A_{i,2}$		$A_{i,3}$		$A_{i,4}$	
	$-\infty$		1.01		1.19		2.13	∞

This illustrates two limitations of set identification. First, preferences are not point identified within the bounds of a set. The choice of community 3 reveals only that the household's preferences lie somewhere in $A_{i,3}$: $1.19 \leq \alpha_i \leq 2.13$. Second, the preference set that corresponds to the highest (lowest) provision of public goods is not bounded from above (below) by the revealed preference logic in (3). These two limitations require that a distribution be specified for α_i . This added information transforms the observed location choices by a population of households into a distribution of preferences.

When vertical differentiation is relaxed, observed location choices are required to set identify more heterogeneous preference parameters. Horizontal differentiation implies households differ in their relative preferences for the two public goods; i.e. $(\gamma_{i,air}, \gamma_{i,school})$ varies across households. This generalization increases the dimensionality of the partition. Figure 1 partitions preference space into regions that rationalize each of the four community choices at *y=$50,000*. The figure illustrates how the identifying power of the indirect utility function differs under vertical and horizontal differentiation. In the vertical case the choice of community 3 indicates that the household's preferences belong to the set: (γ_{air} = .48, $1.19 \leq \alpha \leq 2.13$), which appears in figure 1 as the horizontal line toward the top of the $A_{i,3}$ region. $A_{i,3}$ is the preference set identified by the choice of community 3 in the horizontal case. This comparison illustrates a general principle: *preference sets revealed by vertically differentiated sorting are subsets of their horizontally differentiated counterparts*.

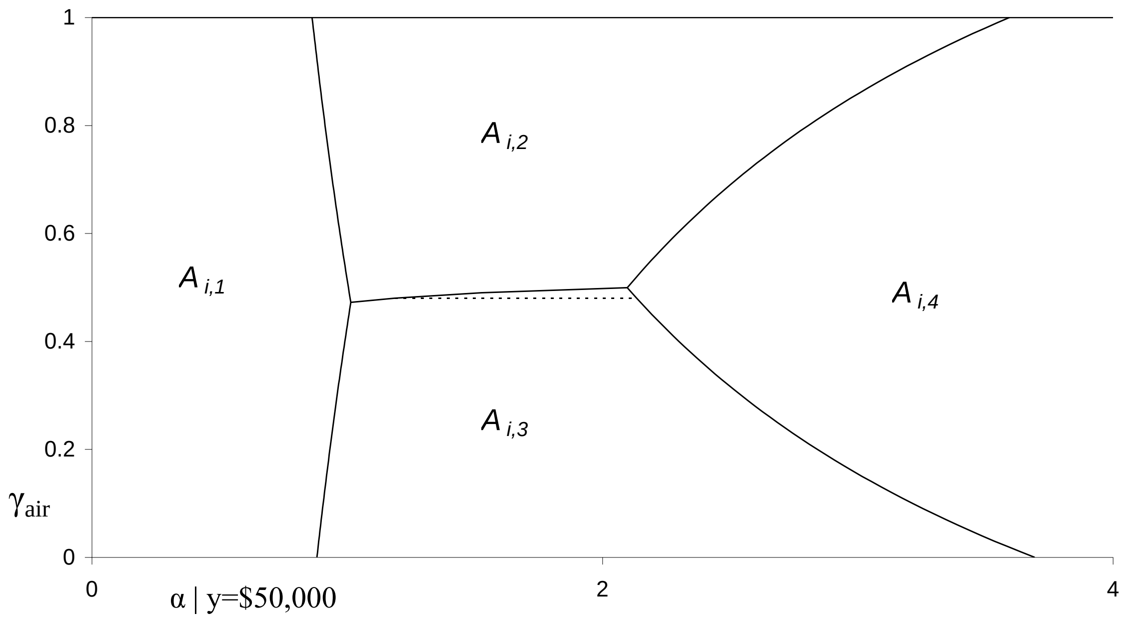


FIGURE 1.—Partitioning Preference Space, Horizontal Differentiation ($\gamma_{school}+\gamma_{air}=1$)

Figure 1 also illustrates how structural restrictions on the utility function control the scope of substitution patterns. With vertical differentiation each community has at most two substitutes, the adjacent communities in the ranking by public goods.[4] With horizontal differentiation the total number of substitutes for each community falls between 2 and J, depending on the number of choices relative to the number of public goods (Anderson, DePalma and Thisse [1992]). The communities that are substitutes will share "borders" in the partition of preference space. Community 2, for example, shares borders with each of the other three communities in figure 1. Consider a marginal increase in the price of housing in community 2. Households that currently reside in 2

[4] The definition of substitution used here is defined as "strong gross substitution" in Anderson, DePalma and Thisse (1992), where k is a substitute for j iff $\partial h_k / \partial P_j > 0$.

but have preferences on the border between 2&4 will respond to the price increase by moving to community 4. Likewise, households on the borders between 2&1 and 2&3 will move to communities 1 and 3. In general, locations that are similar in terms of prices and public goods are more likely to be substitutes than those that are not. Notice that in figure 1 the two communities with intermediate levels of public goods, 2 and 3, share borders with each of the other three locations while the most and least expensive locations, 1 and 4, do not share a border. Because locations 1 and 4 are furthest removed in terms of prices and public goods, it seems natural to expect that there are few, if any, households that consider them to be close substitutes.

2.2. *Introducing the Labor Market*

For working households there are two dimensions of location choice—the choice of a house and the choice of a job. Intuition and recent empirical research suggest these two choices are interrelated (Rhode and Strumpf [2003]). This section expands the theoretical model to allow households with heterogeneous job skills to simultaneously sort among housing communities and labor markets. Under these conditions, the levels of public goods will affect behavior in both markets (Rosen [1979], Roback [1982]). Thus, one might expect job locations to convey additional identifying information about preferences. A single crossing restriction on preferences leads to three properties that must characterize sorting behavior for every household "type" in any locational equilibrium. These properties guarantee that housing and labor market choices convey sufficient information to recover preferences. The primary difference between these properties and the ones derived in Epple and Sieg (1999) arises because a multiplicity of types implies sorting behavior that is less restrictive.

Let the urban landscape be divided into K labor markets that differ in the wage paid to workers of each job skill. With J housing communities and K labor markets, each (j,k) pair represents a unique job-house combination, which will be referred to as a "location" and denoted by $L_{j,k}$. Each location requires a specific commute time, $t_{j,k}$, and households may differ in their aversion to commuting based on their preferences for leisure, ϕ. For a household that commutes between *j* and *k*, let $w_{j,k}(\theta)$ represent wage earnings, where θ describes job skill. Then, a household's income equals $\hat{y} + w_{j,k}(\theta)$, its exogenous non-wage income ($\hat{y}$) plus its wage income.

Utility maximization is similar to (1)-(2), except that households now optimize over two dimensions of location choice and a budget constraint that varies across locations. Equation (4) shows the utility maximization problem for household *i*.

$$(4) \qquad L^*_{j,k} = \max_{j,k} \; V\left[\bar{g}_j(\gamma_i), p_j, \alpha_i, t_{j,k}, \phi_i, y_{i,j,k}\right], \quad \text{where} \quad y_{i,j,k} = \hat{y}_i + w_{j,k}(\theta_i).$$

Holding the community fixed at *j*, a utility-maximizing household will always choose to work in the labor market that provides it with the highest utility, given its job skills. Let $\hat{w}_j(\theta)$ and $\hat{t}_j$ represent the wage and commute time that would maximize utility for a household living in community *j*. Then (4) can be rewritten as (5), with *k* optimized out of the expression.

$$(5) \qquad L^*_j = \max_j \; V\left[\bar{g}_j(\gamma_i), p_j, \alpha_i, \hat{t}_j, \phi_i, y_{i,j}\right], \qquad \text{where} \quad y_{i,j} = \hat{y}_i + \hat{w}_j(\theta_i).$$

For each (γ, ϕ, θ) "type" household, the relevant choice set can be further reduced to a subset of the *J* communities. This is because, conditional on values for γ, ϕ, and θ, some communities may be dominated. A community is dominated if there is another with more public goods and either a sufficiently lower price, a sufficiently higher wage, a sufficiently shorter commute, or some combination of the three. For example, given $\bar{g}_1(\gamma) > \bar{g}_2(\gamma)$, community 1 dominates community 2 if prices and wages are defined such that: $P_1 < P_2$, $\hat{w}_1(\theta) > \hat{w}_2(\theta)$, and $\hat{t}_1 < \hat{t}_2$. No utility-maximizing $(\gamma, \phi, \theta)-$type would ever locate in community 2. Let *R* denote the total number of communities that are not dominated. Then equation (6) shows how the relevant choice set for each $(\gamma, \phi, \theta)-$type relates to the set of all communities in (5), and to the set of all locations in (4).

(6) $$\{L_1,\ldots,L_R \mid \gamma,\phi,\theta\} \subset \{L_1,\ldots,L_J \mid \gamma,\phi,\theta\} \subset \{L_{1,1},\ldots,L_{J,K} \mid \gamma,\phi,\theta\}.$$

Imposing a single crossing restriction on preferences makes it possible to characterize how, in equilibrium, households of each $(\gamma,\phi,\theta)-$type must be sorted across the R communities that are not dominated for that type. Equation (7) shows the slope of an "indirect indifference curve" in $(\overline{g},p)$ space.

(7) $$M[\overline{g}(\gamma), p, \alpha, \hat{y}, t, \phi, w(\theta)] = \left\langle \frac{dp}{d\overline{g}} \middle| V = \overline{V} \right\rangle$$

$$= -\frac{\partial V[\overline{g}(\gamma), p, \alpha, \hat{y}, t, \phi, w(\theta)]/\partial \overline{g}}{\partial V[\overline{g}(\gamma), p, \alpha, \hat{y}, t, \phi, w(\theta)]/\partial p}.$$

Assuming M is monotonically increasing in $(\hat{y} \mid \alpha,\gamma,\phi,\theta)$ and $(\alpha \mid \hat{y},\gamma,\phi,\theta)$, indifference curves in the $(\overline{g},p)$ plane satisfy single crossing in $\hat{y}$ and α conditional on relative preferences, job skills, and the shadow value of time. This restriction has an intuitive interpretation. Roy's Identity implies that $-\partial V(\cdot)/\partial p$ must equal the marginal utility of income, $\lambda = \partial V(\cdot)/\partial y$, times the Marshallian demand for housing, $h[\overline{g}(\gamma), p, \alpha, \hat{y}, t, \phi, w(\theta)]$.

(8) $$M(\cdot) = -\frac{\partial V(\cdot)/\partial \overline{g}}{\partial V(\cdot)/\partial p} = \frac{\partial V(\cdot)/\partial \overline{g}}{\lambda h(\cdot)} = \frac{1}{h(\cdot)}\left[\frac{\partial V(\cdot)/\partial \overline{g}}{\partial V(\cdot)/\partial y}\right].$$

The term in brackets in equation (8) is the Marshallian virtual price of public goods. Therefore, the single crossing restriction implies that the Marshallian virtual price, per unit of housing, is strictly increasing in income and in preferences for public goods relative to private goods.[5]

[5] This property is related to the Willig condition that is often applied together with weak complementarity to identify the Hicksian willingness to pay for changes in public goods. The Willig condition requires the willingness-to-pay per unit of the weak complement to be constant at all levels of income. See Smith and Banzhaf (2004) or Palmquist (2005) for details.

The single crossing property implies that, in equilibrium, three properties characterize sorting by each household type: *boundary indifference, stratification,* and *non-decreasing bundles*.[6] Without loss of generality, let the *R* locations be ordered according to their perceived provision of public goods, $\bar{g}_1(\gamma) < ... < \bar{g}_R(\gamma)$. Boundary indifference requires a household on the "border" between two locations in $(\alpha, \hat{y})$ space to be exactly indifferent between those locations. Equation (9) defines the set of border individuals. It must hold for all $r = 1,..., R-1$.

(9) $$\left\{ (\alpha, \hat{y} \mid \gamma, \phi, \theta): \; V[\bar{g}_r(\gamma), p_r, \alpha, \hat{y}, \hat{t}_r, \phi, \hat{w}_r(\theta)] = V[\bar{g}_{r+1}(\gamma), p_{r+1}, \alpha, \hat{y}, \hat{t}_{r+1}, \phi, \hat{w}_{r+1}(\theta)] \right\}.$$

The non-decreasing bundles property requires that for any two locations in the ordering, $(r, r+1)$ equation (10) must hold.

(10) $\bar{g}_{r+1}(\gamma) > \bar{g}_r(\gamma) \;\Rightarrow$ *at least one of the following must hold*:

$$\text{(i)} \quad p_{r+1} > p_r, \qquad \text{(ii)} \quad \frac{1}{\hat{w}_{r+1}(\theta)} > \frac{1}{\hat{w}_r(\theta)}, \qquad \text{(iii)} \quad \hat{t}_{r+1} > \hat{t}_r.$$

The equation implies that households must "pay" for the additional public goods provided by higher ranked locations through higher housing prices, lower wage income, longer commute times, or some combination of the three. The third property, stratification, requires that households of each type are stratified across the *R* ordered locations by $(\alpha \mid \hat{y})$ and by $(\hat{y} \mid \alpha)$, as defined in (11).

(11) $$\begin{gathered} (\hat{y}_{r-1} \mid \alpha, \gamma, \phi, \theta) < (\hat{y}_r \mid \alpha, \gamma, \phi, \theta) < (\hat{y}_{r+1} \mid \alpha, \gamma, \phi, \theta) \\ \textit{and} \\ (\alpha_{r-1} \mid \hat{y}, \gamma, \phi, \theta) < (\alpha_r \mid \hat{y}, \gamma, \phi, \theta) < (\alpha_{r+1} \mid \hat{y}, \gamma, \phi, \theta) \end{gathered}.$$

---

[6] Boundary indifference and stratification follow from the proof of proposition 1 in Epple and Sieg (1999) because income is separable in non-wage income and wage income. To see why non-decreasing bundles must hold, suppose equation (10) fails for some (*r*,*r+1*) pair. Then *r* must have fewer perceived public goods, more expensive housing, a longer commute, and lower wage income. If so, *r+1* dominates *r*, which implies $r \notin R$, a contradiction.

In the special case where wage income is exogenous to location choice and households are vertically differentiated, the three sorting properties reduce to the ones derived in Epple and Sieg (1999). While the three conditions are necessary for a locational equilibrium to exist, they are not sufficient. Any locational equilibrium must also be characterized by a set of housing prices and wage rates such that no household could increase its utility by changing locations, and all locations are occupied. The estimation strategy in this paper follows Epple and Sieg by recovering values for the preference parameters that justify observed location choices under the assumption that those choices reflect a locational equilibrium.[7]

3. ESTIMATION

3.1. *Indirect Utility Function*

Working households are assumed to possess one of *S* different observable occupations and every household may differ in its preferences ($\alpha_i, \gamma_i, \phi_i, \theta_i$), so households are indexed by both *i* and *s*. Then the indirect utility obtained by household *i,s* in location *j,k* can be expressed as (12).

$$(12) \qquad V_{i,s,j,k} = \left\{ \alpha_i \left(\bar{g}_{i,j} \right)^{\rho} + \left[\exp\left(\frac{\left(\hat{y}_i + w_{i,s,k} \right)^{1-\nu} - 1}{1-\nu} - \phi_i t_{s,j,k} \right) \exp\left(- \frac{\beta P_j^{\eta+1} - 1}{1+\eta} \right) \right]^{\rho} \right\}^{\frac{1}{\rho}},$$

where $\bar{g}_{i,j} = \gamma_{i,1} g_{1,j} + \ldots + \gamma_{i,N-1} g_{N-1,j} + \gamma_{i,N} \xi_j$.

The first term in the CES function represents utility from public goods, and the second represents utility from the private good component of housing and the disutility from commuting. All households are assumed to share the same elasticity of substitution between public and private goods (ρ) as well as the same housing demand parameters:

[7] Epple and Platt (1998) and Sieg et al. (2004) demonstrate existence numerically when income is exogenous and preferences are vertically differentiated.

price elasticity (η), income elasticity (ν), and demand intercept (β). The signs of these parameters provide a test on the consistency of the theoretical model. With $\eta < 0$, $\nu > 0$, and $\beta > 0$, the single crossing restriction implies $\rho < 0$.

Households have horizontally differentiated preferences over a linear index of public goods, $\bar{g}_{i,j}$. Of the *N* public goods in the index, *N-1* are observable. The N^{th} public good ($g_{N,j} = \xi_j$) is not observed by the econometrician.[8] Households differ in the weights they place on each public good in the index $(\gamma_{i,1},....,\gamma_{i,N})$ and in their overall preferences for public goods relative to private goods (α_i). The weights are assumed to sum to 1, allowing α_i to be identified separately as a scaling parameter on the strength of preferences.

As in the theoretical model, a household's income is defined by the sum of its exogenous non-wage income and wage income. The primary earner of each household is assumed to possess skills that qualify them for a certain occupation (e.g. biomedical engineer, locksmith). This is the observable component of job skill indexed by *s*. In the labor market represented by *k*, the average wage for that occupation is $w_{s,k}$. However, a worker's ability to collect that wage if they were to move from their current job depends on (unobserved) idiosyncratic features of their job skill (e.g. education, experience, ability). These features are reflected in a single heterogeneous parameter, θ_i, that represents each worker's labor market mobility *within* their occupation. Equation (13) illustrates how this job skill parameter determines the wage that would be earned at each location.

$$(13) \qquad \begin{aligned} w_{i,s,k} &= w_i, && \text{if the worker occupies } k. \\ w_{i,s,k} &= \theta_i w_{s,k}, && \text{if the worker does not occupy } k. \end{aligned}$$

[8] ξ_j can be interpreted as a composite index of all the *unobserved* public goods under the restriction that they are vertical characteristics; i.e. the weights in the index of unobserved public goods are all constants. This is an example of the "pure characteristics" approach to modeling the utility from a differentiated product (Berry and Pakes [2007]).

At the worker's current job they are observed earning w_i. However, if that worker were to move to a different labor market they may be paid more or less than other workers with the same occupation depending on their relative job skill *within* occupation *s*. For example, if θ_i is greater (less) than 1, the worker would earn more (less) than the average wage for their occupation if they were to move to a new labor market.

The job location choice can also present working households with a long-run tradeoff between leisure time and their consumption of private goods. Holding their home community fixed, a worker may be able to increase their annual wage income by commuting to a more distant labor market. Their willingness to make this commute depends on $\phi_i t_{s,j,k}$, where ϕ_i denotes their preferences for leisure and $t_{s,j,k}$ is the commute time between *j* and *k* for workers in occupation *s*. For a worker with $\phi_i = 0$, there is no disutility from commuting. They would trade a large increase in their commute time for the additional private goods that could be purchased from a marginal increase in their wages. As ϕ_i increases, so does the threshold wage needed to induce the worker to lengthen their commute. By influencing a worker's job location, ϕ_i can affect the amount of income their household has to spend on housing and other private goods.[9]

Working households effectively pay for the public goods provided by their home community through a combination of housing prices, wage rates, and commute times. The job and house location choices they make will reflect their idiosyncratic job skills and their heterogeneous preferences for local public goods and leisure. The richness in this specification for utility poses two key challenges for the inversion process underlying the revealed preference logic of the estimation. It must account for the presence of unobserved public goods and allow for heterogeneity in a subset of the structural parameters. Epple and Sieg (1999), Bayer, McMillan, and Reuben (2005), and Epple, Peress, and Seig (2005) have all developed estimators that address these challenges. However, they each require additional restrictions on the shape of the utility function and assumptions

[9] Some workers are also free to make short-run tradeoffs between leisure time and private goods by varying the number of hours they work. If it were possible to observe their ability to make these tradeoffs, the indirect utility function could be generalized to include hours worked as a third margin for adjustment.

for the distribution of heterogeneous parameters that can be viewed as restrictions on (12). The specification for utility used by Bayer, McMillan, and Reuben would restrict public goods to be perfectly substitutable with the private good component of housing (i.e. $\rho = 1$). Moreover, their estimator treats job locations as fixed so that wage income is exogenous to location choice. Epple and Sieg restrict income to be exogenous to location choice, households to have vertically differentiated preferences for public goods, and the joint distribution of preferences and income to be lognormal. Equation (12) reduces to their specification when ϕ_i and θ_i are dropped along with the i subscripts from $(\gamma_1, \gamma_2, ..., \gamma_N)$, and $f(\alpha, y)$ is lognormal. Epple, Peress, and Seig relax the need for parametric assumptions on the joint distribution of income and preferences, but still require vertical differentiation and exogenous income. Rather than restrict the indirect utility function to satisfy the assumptions needed to implement the existing structural estimators, the remainder of this section develops a new approach.

The new estimator can be decomposed into two stages. The first stage recovers the price of housing in each community ($p_1, ..., p_J$) and the homogeneous housing demand parameters (β, η, ν). These results are treated as known constants during the second stage of the estimation, which simultaneously recovers a composite unobserved public good for each community ($\xi_1, ..., \xi_J$), the homogeneous CES parameter (ρ), and a partition of preference space for the heterogeneous parameters $A(\alpha, \gamma, \phi, \theta)$.

3.2. *First Stage Estimation*

In the theoretical model, housing is treated as a homogeneous commodity that can be consumed in continuous quantities. Under this assumption, the price of housing reflects the cost of consuming the public goods provided by each community. Of course, in practice housing is not homogenous. Its structural characteristics (e.g. bedrooms, bathrooms, sqft.) vary within and between communities, and these differences will be reflected in observable sale prices. This can be addressed if we are prepared to assume that the structural characteristics of housing enter the <u>direct</u> utility function through a sub-function that is homogeneous of degree one and separable from the effect of public goods and the numeraire. Under this restriction, Sieg et al. (2002) demonstrate that the equilibrium locus of housing expenditures defined by a hedonic price function will be separable in the structural characteristics of houses and the effect of public goods, as shown in (14).

$$(14) \qquad e_{j,n} = \bar{h}\left(h_{j,n}\right) \cdot p_j\left(g_{1,j},\ldots,g_{N-1,j},\xi_j\right) \ .$$

The left side of the expression represents expenditures on house *n* in community *j*. The first term on the right side is a "quantity" index of housing that depends on a vector of structural characteristics ($h_{j,n}$). By condensing all the information about the structural characteristics of a house into a single number, the index provides an empirical analog to the concept of a homogeneous unit of housing from the theoretical model. The second term represents the price of a homogeneous unit of housing in community *j*, which depends on the public goods it provides, observed and unobserved. Taking logs of (14) produces the housing price hedonic model in (15), where $\mu_{j,n}$ represents measurement error.

$$(15) \qquad \ln\left(e_{j,n}\right) = \ln\left[\bar{h}\left(h_{j,n}\right)\right] + \ln\left[p_j\left(g_{1,j},\ldots,g_{N-1,j},\xi_j\right)\right] + \mu_{j,n} \ .$$

Given a parametric form for (15) and data on housing transaction prices and their structural characteristics, the price of housing in each community can be recovered as a community-specific fixed effect.

Estimates for the price of housing can be used along with data on housing expenditures and household income to recover the homogenous housing demand parameters (β,η,ν). Using Roy's Identity, an individual household's demand for housing can be derived from the indirect utility function as equation (16).

$$(16) \qquad \bar{h}_i = \beta p_z^{\eta} y_i^{\nu} .$$

Multiplying both sides of (16) by the price of housing and taking logs produces the expression for housing expenditures in (17), where expenditures are assumed to be measured with error (ε_j). The N superscript indicates that the expression is evaluated for a household at the Nth quantile in the income distribution for community *j*. The intercept in the demand for housing can be estimated together with the price and income elasticities by regressing quantiles of the distribution of annualized housing expenditures, e_j^N, on the price of housing and quantiles of the income distribution, y_j^N. While a single quantile is sufficient to identify the demand parameters, adding data on additional quantiles can increase the efficiency of the estimation.

$$\ln\left(e_j^N\right) = \ln(\beta) + (\eta + 1)\ln\left(p_j\right) + \nu \ln\left(y_j^N\right) + \varepsilon_j . \tag{17}$$

Since housing prices were estimated as fixed effects in a hedonic regression of (15), they will be measured with error. The observable public goods can be used as instruments for price to address the potential endogeneity problem. In addition, non-wage income can be used as an instrument for income, which will be endogenous if a worker's wage income depends on their residential location choice. Assuming the error terms in (17) are uncorrelated across different quantiles of the distribution of income and expenditures, the quantiles can be stacked and the regression can be run using 2SLS.[10]

Throughout the second stage of the estimation the first stage estimates are treated as known constants.[11] To reduce notation in the following discussion, let δ represent the first stage results plus all the data on attributes of locations: $\delta = \left[\beta, \eta, \nu\ ;\ \ p, g, w, t\right]$.

3.3. *Second Stage Estimation*

The estimator uses an iterative process to simultaneously recover all the second-stage parameters. The iterative structure is based on solving for a point estimate of ρ. On the first iteration, a starting value (ρ^0) is used to solve for a vector of unobserved public

[10] Alternatively, if the error terms are expected to be correlated across quantiles, the estimation could be performed using GMM or using SUR with restrictions on the parameters across equations.
[11] Alternatively, endpoints of a confidence interval on each parameter in (17) could be used to place bounds on the second stage parameters.

goods ($\xi_1^0,...,\xi_J^0$) which are then used together with ρ^0 to partition preference space. The resulting partition, $A^0(\alpha,\gamma,\phi,\theta)$, is used to evaluate an objective function that equals zero at the true value of ρ. Then, the value of the objective function is used to choose a new value for the CES parameter (ρ^1) to be used during the second iteration. This process terminates when additional changes in ρ do not lead to further improvements in the objective function. The remainder of this subsection first describes how $\xi_1,...,\xi_J$ and $A(\alpha,\gamma,\phi,\theta)$ are identified conditional on a value for ρ and then describes the objective function used to identify ρ.

If unobserved public goods influence households' location choices, they should also influence the price of housing. Under the maintained assumption that households have nonnegative preferences for public goods, the price of housing will be strictly increasing in unobserved public goods as in (18).[12]

(18) $$\frac{\partial p(g_1,...,g_{N-1},\xi)}{\partial \xi} > 0 \ .$$

If (18) holds, the price of housing in each community that was recovered as a fixed effect in (15) should contain information about the provision of public goods in that community. More precisely, after controlling for the variation in the price index due to observed public goods, the remaining variation can be attributed to unobserved public goods. However, theory does not suggest a functional form for the relationship between p and $g_1,...,g_{N-1},\xi$. Furthermore, ξ may be correlated with the observed public goods. For example, it has been suggested in the literature that public school quality will be correlated with ξ if unobserved public goods influence the location choices made by households with high incomes or strong preferences for school quality (e.g. Epple and Sieg [1999], Bayer et al. [2007]). Given the potential for endogeneity and functional form indetermi-

[12] Bajari and Benkard (2005) prove a hedonic price function exists and is strictly increasing in ξ if utility satisfies differentiability, continuity, and nonsatiation in ξ and the numeraire. These conditions are satisfied for (12).

nacy, the strategy used here is to recover ξ from the observable data using the control function approach to instrumental variables estimation of a nonseparable function (Chesher [2003], Imbens and Newey [2008]).

Suppose one of the public goods, g_1, is correlated with ξ but a vector of instruments $\tilde{Z}$ exists so that the reduced form for g_1 can be written as $g_1 = g_1(Z, \psi)$, where ψ is a scalar disturbance that reflects the endogenous component of g_1, and $Z = \left[\tilde{Z}, g_2, ..., g_{N-1}\right]$. The vector of unobserved public goods can be recovered from data on p, g, and Z if two additional assumptions are satisfied:

(19.a) $\dfrac{\partial g_1(Z, \psi)}{\partial \psi} > 0$. (19.b) $Z \perp \psi, \xi$.

When (19.a) and (19.b) hold, Imbens and Newey (2008) demonstrate that g_1 will be independent of ξ after conditioning on c, a control variable defined by the quantiles of the conditional distribution of g_1:

$$c_j = F_\psi(\psi_j) = F_{g_1|Z=Z_j}(g_{1j}). \tag{20}$$

After recovering the control variable, the quantiles of the distribution of the unobserved public good can be recovered by integrating over the conditional distribution of prices with respect to the control variable. This result is shown as (21) using $p = f(g, c, \xi)$ to represent the nonparametric mapping between community housing prices and public goods.

$$F_\xi(\xi_j) = \int F_{p|g,c=g_j,c_j}(p_j)\, dc = \int F_{p|g,c=g_j,c_j}\left[f(g_j, c_j, \xi_j)\right] dc \ . \tag{21}$$

A variety of nonparametric methods can be used to estimate (20) and (21). Regardless of the method used, the estimated quantiles in (21) represent a monotonic transformation of the unobserved characteristic itself since, assuming ξ has a continuous distribution, it can be normalized such that its marginal distribution is U[0,1]. This normalization implies $\xi = F_{\xi}(\xi_j)$.

The quantile IV methods developed by Chernozhukov and Hansen (2005) and Chernozhukov, Imbens, and Newey (2007) offer a potential strategy for addressing the more general case of multiple endogenously determined public goods. At the other extreme, if all the public goods can be treated as exogenous the process of recovering ξ simplifies to estimating the quantiles of the distribution of prices, conditional on observed public goods. Assuming observed and unobserved public goods are independently determined, the results from Matzkin (2003) imply that the quantiles of the distribution of the unobserved public good will equal the quantiles of the price distribution, conditional on observed public goods.

Importantly, the estimated values of ξ and ρ must permit the indirect utility function to explain every observed location choice. In other words, each location must maximize utility for some set of values for the heterogeneous parameters. This requires a certain degree of smoothness in the relationship between the price of housing and the unobserved public good.[13] In practice, the minimum bandwidth that provides this smoothness may exceed the bandwidth that would otherwise be chosen to address the bias/efficiency tradeoff from estimating (21). In the estimation, this is treated as a constraint on the bandwidth. The estimator starts with an approximation to the optimal bandwidth. Then, if necessary, the bandwidth is increased until the estimator finds values for the heterogeneous parameters that justify every observed location choice.

Given δ, $\xi_1,...,\xi_J$, and a value for ρ, location choices can be expressed as a function of preferences for public goods, the opportunity cost of time, and unobserved job skill. The partitioning process inverts this relationship, using the logic of revealed preferences to recover values for the heterogeneous parameters that rationalize observed location choices. This step of the estimation manifests Tiebout's logic that location

[13] This is a common feature of pure characteristics-based models such as Feenstra and Levinsohn (1995), Epple, Peress, and Sieg (2005), and Bajari and Benkard (2005). Alternatively, in applications of the random parameters logit model such as Berry, Levinsohn, and Pakes (1995) and Bayer, McMillan, and Reuben (2005) the idiosyncratic logit error terms "pick up the slack" in explaining choices.

choices reveal preferences.

The borders that delineate the partition of preference space are implicitly defined by the system of equations that arise from applying the boundary indifference condition in equation (9) to the indirect utility function in (12). This system is nonlinear. Consequently, the borders cannot be expressed analytically and when preference space exceeds two dimensions it is infeasible to solve for them numerically. Instead, the estimator recovers an approximation to the partition of preference space by sampling over it uniformly.[14]

The sampling is done by a Gibbs algorithm that takes a large number of uniform draws from each region of the partition. For example, suppose we want to sample uniformly over region $A_{i,3}$ of the partition in figure 2. To start the Gibbs sampler, one must first locate a point somewhere in $A_{i,3}$. In the figure, the starting value is denoted by $*_0$. The first step is to condition on all but one coordinate and solve for bounds on the remaining coordinate. In the figure, this is done by conditioning on γ_{air} and solving for the bounds on α, which are *0.96* and *2.55*. Use these bounds to take a random uniform draw. Suppose the result is $\alpha = 2.3$. From here, condition on $\alpha = 2.3$, solve for the bounds in the γ_{air} dimension, and take a random uniform draw on γ_{air}. In the figure, the new bounds are *0.0* and *0.4*, and the new uniform draw is *0.15*. Together, the two conditional uniform draws (*2.3, 0.15*) define the first unconditional draw from the region, $*_1$. This process can be repeated, using $*_1$ to find $*_2$ and so on. The result is a randomly chosen uniform distribution of points within $A_{i,3}$ that approximates its shape.[15]

Operationally, the process of partitioning preference space relies on the three conditions used to characterize sorting behavior in the theoretical model. *Non-decreasing bundles* identifies locations that have adjacent regions in the partition. *Boundary indifference* defines the borders that delineate those regions, and *stratification* guarantees that

[14] Similar methods have been used in different empirical contexts under the simplifying assumption of a linear utility function (Feenstra and Levinsohn [1995]; Bajari and Benkard [2005]).
[15] See Geweke (1996) for a general description of Gibbs sampling and see the supplemental appendix to this paper for additional computational details.

each region is connected in $(\alpha \mid \hat{y}, \gamma, \phi, \theta)$.

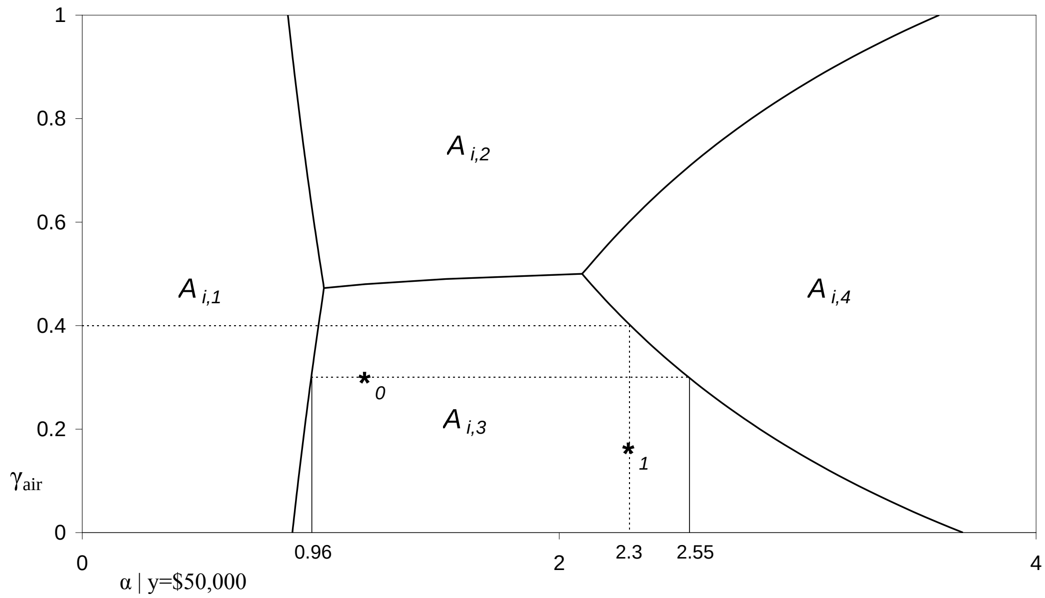


FIGURE 2.—Using the Gibbs Sampling Algorithm to Partition Preference Space

While observed location choices are sufficient to identify $\xi_1,...,\xi_J \mid \rho$, and $A(\alpha, \gamma, \phi, \theta) \mid \rho$, they are not sufficient to separately identify ρ, $\xi_1,...,\xi_J$, and $A(\alpha, \gamma, \phi, \theta)$ without some prior knowledge of the relationship between preferences and income. This information can be supplied by specifying a parametric form for their joint distribution (Epple and Sieg [1999]) or assuming they are independent for a subset of households (Epple, Peress, and Sieg [2005]). The later approach is illustrated here.

All else constant, the interaction between ρ and $\hat{y}$ in the CES indirect utility function dictates how income shocks affect the desired bundle of housing and public goods. This relationship can be inverted to identify ρ from the location choices made by households that are identical except for their non-wage income. Put differently, the observed stratification by income of (otherwise) identical households reveals the extent to which they substitute public goods with the private good component of housing.

Let $F_s(\alpha, \gamma, \phi, \theta)$ denote the distribution of the heterogeneous parameters for a subset of households, s, for which $F_s \perp \hat{y}$. Suppose this subset can be further divided into two groups with non-wage income $\hat{y}_1$ and $\hat{y}_2$. Sampling over the corresponding partitions will produce two approximations to F_s: $\tilde{F}_{s,1}$ and $\tilde{F}_{s,2}$. These conditional distributions will equal the unconditional distribution when the partitioning process is performed at the true value of the CES parameter, $\rho = \rho_0$, as depicted in equation (22).

$$(22) \qquad F_s(\alpha,\gamma,\phi,\theta) = \tilde{F}_{s,1}(\alpha,\gamma,\phi,\theta \mid \hat{y}_1,\rho,\delta,\xi) = \tilde{F}_{s,2}(\alpha,\gamma,\phi,\theta \mid \hat{y}_2,\rho,\delta,\xi), \quad \text{for } \hat{y}_1 \neq \hat{y}_2.$$

The equalities will not hold for other values of the CES parameter. This follows from the observation that the boundary indifference loci defining the partition of preference space are nonseparable in $(\rho,\hat{y})$. A movement in ρ away from its true value will distort the boundaries of the partition to a different extent for $\hat{y}_1$ and $\hat{y}_2$, leading to predictions for $F_{s,1}$ and $F_{s,2}$ that differ from the true distribution and from each other: $\tilde{F}_{s,1} \neq \tilde{F}_{s,2} \neq F_s$. The estimator applies this logic to recover the value of ρ that minimizes the predicted difference between $F_{s,1}$ and $F_{s,2}$, as shown in equation (23).

$$(23) \qquad \min_{\rho} \left\| \tilde{F}_{s,1}(\alpha,\gamma,\phi,\theta \mid \hat{y}_1,\rho,\delta,\xi) - \tilde{F}_{s,2}(\alpha,\gamma,\phi,\theta \mid \hat{y}_2,\rho,\delta,\xi) \right\|.$$

This equation provides a general expression for the objective function that forms the basis for the second stage of the estimation. If location choices can be observed for *s*-type households at more than two income levels, the efficiency of the estimation may be improved by minimizing the difference between the predicted distributions for all pairwise combinations of income. In general, evaluating the objective function requires partitioning preference space at each of the $d = 1,\ldots,D$ income levels and then sampling from those partitions to obtain $\tilde{F}_{s,1},\ldots,\tilde{F}_{s,D}$. This process must be repeated, updating ρ on each step, until the relevant convergence criteria are satisfied.

3.4. *Using* $A(\alpha, \gamma, \phi, \theta)$ *to Recover the Distribution of Preferences*

Translating $A(\alpha, \gamma, \phi, \theta)$ into a distribution of preferences requires taking draws from each region of the partition in proportion to the number of households who live and work in the corresponding locations. Draws can be taken to exactly match the share of residents predicted to live in each location with the actual share observed in the data. However, the values for the heterogeneous preference parameters which underlie these predicted shares need not be unique.

To see how more than one distribution of preferences may be capable of matching predicted and observed population shares, consider the two points in figure 2: $*_0$ and $*_1$. Both points are within the set of values for α, γ that make community 3 the utility-maximizing choice for retired households with income of $50,000. Suppose we observe 750 of these households living in community 3. One could match predicted and observed population shares by taking 750 draws on $*_0$, by taking 750 draws on $*_1$, or by taking 750 draws from anywhere within $A_{i,3}$. Each of these alternatives will correspond to a different distribution of preferences, holding constant the set of draws taken from all other regions of the partition.

The lack of uniqueness in the distribution of preferences can be interpreted two ways. First, if each household's true preferences can be represented by a single point in the partition, the lack of uniqueness simply underscores the fact that if all of the choice-specific characteristics are discrete, some distributional assumptions are needed to point-identify the household's preferences from cross-section data.[16] Alternatively, the lack of uniqueness can be interpreted as a mild relaxation of the restriction that households must have stable preferences in a locational equilibrium. Given all of the attributes of job and house locations which serve as features of a dual-market locational equilibrium, a household's preferences may vary *within* the region of the partition that corresponds to their observed location choice without inducing them to move. This follows from the revealed

[16] The distributional assumptions need not be parametric restrictions on the shape of the distribution. They may consist of independence assumptions on the joint distribution of observed and unobserved variables (for example, see Briesh et al [2007]).

preference logic in (3) which implicitly defines each region of the partition.[17]

Since there is a one-to-one mapping between preferences and welfare measures, the lack of uniqueness in the distribution of preferences means there may be multiple, observationally equivalent, distributions of welfare measures for a given shock to the equilibrium. Fortunately, it is possible to identify bounds on the range of welfare measures that would be consistent with each household's observed behavior. For example, let $*^{L}_{j,k}\big|y$ and $*^{H}_{j,k}\big|y$ denote the points in the partition that correspond to the lowest and the highest willingness-to-pay that would be consistent with the choice to live in location j,k, for a household with income y. Then, conditional on income, partial equilibrium welfare measures calculated at these two points bound the range of plausible estimates for households living in j,k. Repeating this process for each household generates upper and lower bounds on the distribution of welfare measures for a given shock. The empirical magnitude of bounds on the willingness-to-pay for improved air quality is investigated as part of the application to Northern California.

4. DATA

The model was estimated using data from Northern California's two largest population centers: the San Francisco and Sacramento Consolidated Metropolitan Statistical Areas (CMSA). Together, the two CMSAs contain about 9 million people, roughly 25% of the state's population and 3% of the U.S. population. The region is largely self-contained. Only 1.5% of its workforce commutes to a job outside the region. While the two CMSAs are adjacent, their major business districts are 80 to 120 miles apart—far enough to limit commuting, but close enough that most households could move from one to the other without alienating family and friends, or having to readjust to a dramatically different environment. The closeness between the regions is also apparent in data on recent movers. Between 1995 and 2000, San Francisco was the top destination for households moving out of Sacramento. Likewise, San Francisco was the top origin of households that moved

[17] If preferences are unstable, each region of the partition defines the set of values for the heterogeneous preference parameters that satisfy the "strictly unambiguous" revealed preference relation defined by Bernheim and Rangel (2009).

into Sacramento. Together with physical proximity, these migration patterns suggest it is reasonable to treat both CMSAs as part of the same locational choice set. [18]

The data were generated in three steps. First, the study region was divided into housing communities and work destinations, and the observable component of job skill was defined. Second, the set of all possible job-house combinations was reduced to a set of admissible locations, and for each of these the distribution of non-wage income by occupation was obtained. Finally, data were obtained for a set of characteristics that differentiate communities and jobs. Each step is briefly described before proceeding to the estimation results, with additional details provided in a supplemental appendix.

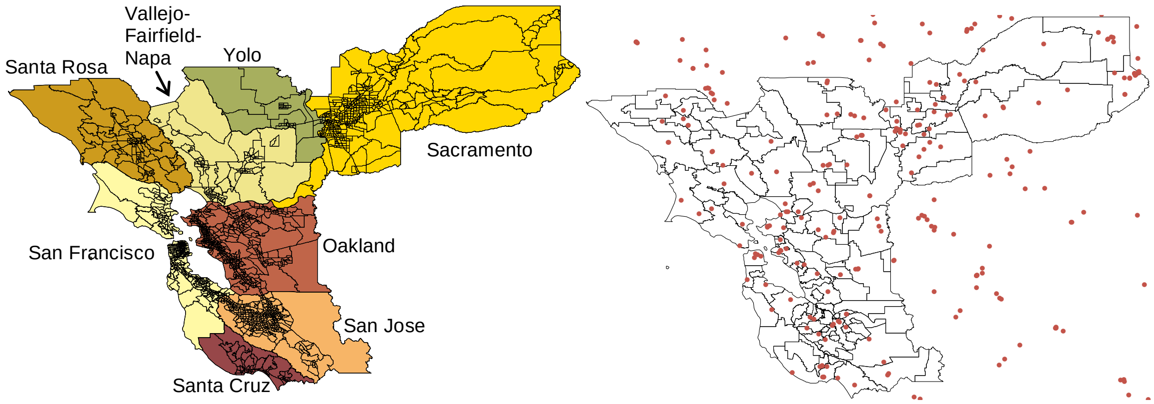
As in most sorting applications, housing communities are defined as unified school districts. Exceptions are made for primary and secondary districts that do not belong to a unified district, and for the city of San Francisco which was divided into 11 supervisorial districts.[19] The resulting housing component of the choice set contains 122 communities. Work destinations are defined as Primary Metropolitan Statistical Areas (PMSA), which resemble distinct labor markets.[20] Figure 3A shows how the region is divided into eight PMSAs and the density of Census tracts (overlaid on figure 3A) illustrates that the population is mostly concentrated around the San Francisco Bay and the city of Sacramento. Finally, a household's job skills are classified according to the occupational category of its primary earner, using the 22 occupational categories in the Standard Occupational Classification System (e.g. *managers*, *healthcare support workers*, etc.).[21] All retired households comprise an additional category.

[18] California's property tax code presents a potential caveat to the notion that households can be treated as freely mobile. To the extent that market values increase by more than 2% annually, households are effectively taxed for moving between two homes of comparable value. However, past evidence has suggested that the effect on mobility is quite small. O'Sullivan et al (1995) report that with average annual inflation of 13%, Proposition 13 increases the time between moves by only 2 months for the average household. Additional detail is provided in section 2.1 of the supplemental appendix.

[19] All public schools in San Francisco are incorporated into a single school district, which comprises 10% of the total population in the study area.

[20] The Census Bureau describes a PMSA as "a large urbanized county or cluster of counties…that demonstrate very strong internal economic and social links, in addition to close ties to other portions of the larger area [CMSA]".

[21] 60% of married couples in the study region reported both the husband and wife working in 1999. While a dual-earner job search would be an interesting extension, it is not possible given present data limitations.



A. Primary Metro Areas & Census Tracts *B. School Districts & Monitoring Stations*

FIGURE 3.—The Regional Landscape

The set of all possible community-PMSA combinations was reduced to 268 admissible locations which comprise the choice set used to estimate the model. The criterion used to define an admissible location is that it must account for at least 500 working households (0.02% of the working population). This rule effectively excludes multiple-hour commutes between opposite ends of the study region, and most commuting between the two CMSAs. 99% of working households live and work in the 268 admissible locations. For each of these locations, distributions of non-wage income by occupation were generated from publicly available special tabulations of Census data. For each community, data were collected on the price of housing and the provision of two public goods, air quality and school quality. Then for all the admissible work locations associated with each community, data were collected on the mean wage rate and mean commuting time for workers in each occupational category.

Data on individual housing transactions were compiled from records in the Assessor's office of each county and contain the price and characteristics of most houses sold in the region between 1995 and 2005. These data were filtered to eliminate observations with apparent errors, those lacking information on structural characteristics, nonresidential properties, and outliers—specifically the most expensive and least expensive 0.5% of sales. The resulting data set contains 540,642 housing transactions which were converted into annual rents using the formula suggested by Poterba (1992).

Ozone concentrations are used as a proxy for air quality. Ozone is an attractive proxy because it is the chief component of urban smog which, for households, is perhaps the most readily observable measure of air quality. The California Air Resources Board records hourly concentrations of ozone at monitoring stations throughout the state. Figure 3B overlays the location of 210 monitoring stations on school districts in the study region. The ozone measure used in this analysis is the average of the top 30 1-hour daily maximum readings (in parts per million) recorded at each monitoring station during the course of a year. Households are assumed to be primarily concerned with air quality near their home, not their job. Under this assumption, community-specific measures are constructed by first assigning to each house the ozone measure recorded at the nearest monitoring station, and then taking an average over all the houses in the community. Then, to control for annual fluctuation in ozone levels, the process was repeated for 1999, 2000, and 2001, and the results averaged. The final measure ranges from 0.031 in the highest air quality community to 0.106 in the lowest.

TABLE I
Descriptive Statistics for 122 Housing Communities

Observed Attribute	Source	Mean	St. Dev.	Min	Max
Community Size (population share)	Census	0.008	0.008	5.45E-05	0.047
Ozone (parts per million)	CA Air Resources Board	0.069	0.015	0.031	0.106
Academic Performance Index	CA Dept. of Education	706	93	528	941
Household Total Income (25th quantile)	Census	36,241	12,195	10,548	77,705
Household Total Income (50th quantile)	Census	65,112	21,195	27,446	147,630
Household Total Income (75th quantile)	Census	105,580	33,698	55,155	226,330
Annual Housing Expenditures (25th quantile)	*Dataquick*	27,825	12,565	9,156	88,082
Annual Housing Expenditures (50th quantile)	*Dataquick*	37,275	16,240	12,166	100,280
Annual Housing Expenditures (75th quantile)	*Dataquick*	48,345	21,127	16,407	123,620

Data on school quality come from the California Department of Education. The measure used in this study is the Academic Performance Index (API), a composite index of standardized test scores, weighted across all subjects and grade levels. For each community in the study region, a three-year average API was constructed by weighting the score of each school in the community by its number of students from 1999-2001. The resulting measure ranges from 528 to 941. Table I reports summary statistics for the API and other community characteristics.

For each occupational category and PMSA, mean annual wages were obtained from the California Employment Development Department.[22] Wages can very substantially between PMSAs, even for aggregate job categories. Workers with jobs in the *construction and excavation* category are paid 32% more in San Jose than in Sacramento, for example. Some of this variation may reflect local cost-of-living adjustments in markets where housing is particularly expensive, like San Jose and San Francisco. The variation may also reflect unobserved heterogeneity in the mix of jobs within each category, or location-specific attributes of jobs.

Finally, data on the mean time for every tract-to-tract commute were taken from the Census Transportation Planning Package special tabulation. These figures were used to calculate an average travel time between each home community and PMSA, weighted by the share of workers making each tract-to-tract commute. The resulting average one-way commute time ranges from 1 to 114 minutes, with a mean of 36 minutes and a standard deviation of 19 minutes. Traffic is a major contributor to the relatively high average commute time. Most workers (82%) live and work in the same PMSA.

5. RESULTS

This section compares the results from implementing the new “dual-market” estimator to the results from two special cases—the Epple-Sieg model, and an intermediate version of the model that admits horizontal differentiation but treats wage income as exogenous. In the dual-market case, the choice set consists of the 268 (housing community, labor market) combinations and a household’s income varies across locations depending on the oc-

[22] Wages include base pay, production bonuses, tips, and cost-of-living adjustments, but exclude nonproduction bonuses, overtime pay and the value of benefits.

cupation of its primary earner and the required commute time. This framework nests the other two versions of the model as special cases. The intermediate "single-market horizontal" model restricts income to be exogenous so that households only choose among the 122 communities. Finally, the "single-market vertical" case restricts income to be exogenous, preferences to be vertically differentiated, and the joint distribution of income and preferences to be lognormal. This is the Epple-Sieg model. The three models can be formally related in terms of equation (12), the dual-market indirect utility function. The single-market horizontal version of the model drops the job skill and time cost parameters (θ_i, ϕ_i) so that total household income is invariant to location choice. This same restriction is imposed in the vertical model, which also assumes $f(\alpha, y)$~lognormal and drops the *i* subscripts from $(\gamma_1, \gamma_2, ..., \gamma_N)$.

Since the differences between the three versions of the model do not affect $p_1,..., p_J$ and β, η, ν, the first stage of the estimation was only performed once. Similarly, the second stage of the estimation was performed simultaneously for the two horizontal models; in other words the same estimates for ρ and $\xi_1,...,\xi_{122}$ were used to recover an approximation to the partition preference space for the single and dual-market models. The only difference is that two additional dimensions of preference space were partitioned in the dual-market case (θ, ϕ). This isolates the way that including job opportunities in the model affects the resulting partition of preference space. Meanwhile, in the single-market vertical case, the second-stage parameters were estimated using the GMM approach developed by Sieg et al. (2004). Comparing the results to those from the two horizontally differentiated models provides the means to evaluate the economic implications of introducing horizontally differentiated preferences and job opportunities into the Epple-Sieg sorting framework, while simultaneously relaxing their lognormal assumption on the joint distribution of income and preferences.[23]

[23] The combined estimation time for all three models is 46 hours on a Dell 690 precision workstation. Virtually all of this time is spent partitioning preference space. While the computational burden is substantial, it does not preclude estimating the model for larger geographic areas. Importantly, computational time depends linearly on the number of choices. This follows from the discrete-continuous formulation of the Gibbs algorithm described in section 1 of the supplemental appendix.

Finally, instrumental variables were used to address the potential endogeneity of school quality during the estimation of $\xi_1,...,\xi_{122}$. The concern is that peer effects in the production of education may interact with the sorting process to induce correlation between the composite unobserved public good and the observable variation in test scores. For example, the production of school quality is often modeled as a function of community income (e.g. Calabrese et al. [2006]; Ferreyra [2007]).[24] If higher income households tend to locate in communities with higher values for ξ, these same communities will tend to have higher quality public schools. This was addressed by developing instruments for school quality based on the income ranking of communities, as in Epple and Sieg (1999). The validity of this approach relies on the assumption that unobserved public goods do not affect the overall ranking of communities by income, though they may influence the location decisions of some households. Given this assumption, the income ranks were used to develop moment conditions for GMM estimation of the vertical model (Sieg et al. [2004]) and were also used as instruments for the control function approach to estimating $\xi_1,...,\xi_{122}$ in the horizontal model.

5.1. *First Stage Estimation Results*

In the first stage of the estimation, the 540,642 observations on individual real estate transactions were used along with income distributions for each community to estimate an index of housing prices and the homogeneous housing demand parameters. First, equation (15) was estimated by regressing the sale price of a home on the number of bedrooms, number of bathrooms, lot sizes, building sizes, age of each house, a dummy variable for condominiums, and a set of community-specific fixed effects.[25] The community-specific fixed effects recovered from the regression indicate that housing in the most expensive community costs 6.5 times as much as in the cheapest community. After nor-

[24] Both models also depict homeowners voting on the property tax rates used to finance public education. This form of endogeneity is less relevant in California because the state's school finance system equalizes expenditures per student across districts.

[25] All continuous variables were measured in logarithms. The regression, which also included interactions of the dependent variables, had an R^2 of 0.81. Complete results are reported in the supplemental appendix.

malizing by the lowest price, the index ranges from 1.00 in Sacramento's Grant Union high school district to 6.51 in San Francisco's second supervisorial district.[26] Overall, the distribution is consistent with the conventional wisdom that the Bay Area is an expensive place to live. The 11 cheapest communities are all located in the Sacramento PMSA, while 24 of the 25 most expensive communities are in the San Francisco and San Jose PMSAs. Despite the spatial concentration of communities with extreme values for the price index, there is considerable variation within most PMSAs. The price of housing varies by more than 100% between the most expensive and least expensive communities in Oakland, San Francisco, San Jose, and Vallejo. Furthermore, these ranges overlap for 21 of the 29 possible PMSA pairings.

TABLE II
Housing Demand Parameter Estimates (standard errors)

Specification	Demand Constant (β)	Price Elasticity (η)	Income Elasticity (ν)	R^2
OLS	29.72 (1.23)	-0.33 (0.02)	0.58 (0.02)	0.888
IV: price = f(ozone, score) income = f(nonwage income)	11.97 (1.59)	-0.38 (0.03)	0.66 (0.04)	0.878

The housing price index was used together with data on the distribution of income and housing expenditures in each community to estimate the demand for the private good component of housing. Specifically, equation (17) was estimated by regressing quantiles from the distribution of annualized housing expenditures in each community on the price of housing and quantiles from the income distribution for working households. The 25th, 50th, and 75th quantiles were used. As discussed earlier, there is reason to expect both prices and income may be endogenous. Therefore the expenditure function was estimated using 2SLS in addition to OLS. The 2SLS regression used the observed public goods

[26] San Francisco's 2nd supervisorial district comprises the area just southeast of the Golden Gate Bridge, including the city's affluent Marina district.

as instruments for the price of housing and the 25th, 50th, and 75th quantiles from the distribution of non-wage income as instruments for total income. Table II reports the results.

Including instruments in the regression produces a modest increase in the income elasticity and a modest decrease in the price elasticity relative to OLS. As the elasticities increase in absolute magnitude the demand intercept decreases. The estimates for the price elasticity are similar to the results from previous sorting applications. For example, the 2SLS estimate $(\hat{\eta} = -0.38)$ falls near the middle of the range reported in the existing literature (-0.01 to -0.70).[27] While the corresponding estimate for the income elasticity $(\hat{\nu} = 0.66)$ falls slightly below the range of results from previous studies (0.73 to 0.94), their 95% confidence intervals overlap. The p-values from Basmann, Hansen, and Sargan overidentification tests range from 0.241 to 0.276.

5.2. *Second Stage Estimation Results: Single-Market Vertical Model*

If wage income is exogenous, households have identical relative preferences for different public goods, and the shape of the joint distribution of income and preferences is known to be lognormal, then all the remaining structural parameters can be estimated simultaneously using the GMM estimator developed by Sieg et al. (2004). Table III reports the results from using their estimator to recover the CES parameter, the parameters that characterize the joint lognormal distribution of income and preferences, and the constant weight on air quality in the public goods index.[28]

[27] This includes all sorting applications that have estimated (17) directly or included it as a moment condition in GMM estimation: Epple and Sieg (1999), Walsh (2007), Sieg et al. (2004), and Epple et al. (2005). Polinsky (1977) reports a lower range of estimates (-0.87 to -0.67) in his summary of consistent micro models. However, unlike the sorting literature, these earlier studies did not control for variation in the structural characteristics of homes.

[28] The residual to one of the moment conditions defines the composite unobserved public good in each community. The estimation process also recovers the overall level of public goods provision in the cheapest community as an incidental parameter. Its estimated value was 0.273 (0.455). Following Seig et al., the weight on school quality was normalized to one.

TABLE III
Second-Stage Parameter Estimates: Single-Market Vertical Model

mean ln(y)	standard deviation ln(y)	mean ln(α)	standard deviation ln(α)	corr(y,α)	CES parameter	weight on air quality
μ^y	σ^y	μ^α	σ^α	λ	ρ	γ_{air}
11.057 (0.006)	1.309 (0.001)	1.132 (0.538)	0.366 (0.001)	-0.539 (0.001)	-0.013 (0.000)	0.149 (1.131)

Most of the parameters in table III are precisely estimated and similar in magnitude to the results in Sieg et al. The negative correlation between income and preferences for public goods ($\lambda < 0$) reflects the fact that there is considerable overlap in the community-specific income distributions. Alternatively, if λ were positive, the model would predict almost no overlap in the range of income within different communities. The negative value for ρ indicates the elasticity of substitution between public and private goods is less than one, which implies the marginal willingness-to-pay for public goods is increasing in income. This is consistent with the single-crossing restriction on preferences, providing a consistency check on the theoretical model.

Recall that in a vertically differentiated model households can be ordered along an interval according to their preferences for public goods. Figure 4 illustrates part of the implied ordering for households with income equal to \$67,500:

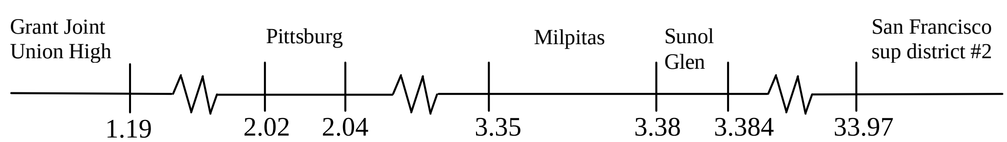


FIGURE 4.—Preference Regions, Single-Market Vertical Case

These results imply that a household with $\alpha < 1.19$ and an annual income of \$67,500 will maximize its utility by purchasing housing in Grant Joint Union high whereas a household with the same income and $\alpha > 33.97$ will purchase housing in San Francisco's second supervisorial district.

The positive value for γ_{air} in table III indicates that, all else held constant, households with higher values for α will be willing to pay more for a small improvement in air quality. However, the point estimate for γ_{air} is not precisely estimated. The horizontal model relaxes the restriction that households share identical values for this parameter.

5.3. *Second Stage Estimation Results: Single-Market Horizontal Model*

Implementing the horizontal estimator requires identifying a subset of households for whom preferences and income are independent. Using only those households, the (iterative) estimation can be performed to obtain consistent estimates for ρ, ξ, and an approximation to the partition of preference space that rationalizes the location choices made by those households. Then, treating the estimates for ρ and ξ as known constants, preference space can be partitioned once for the remaining households. This strategy was used to recover ρ and ξ from data on retired households.

Retired households were a strategic choice for two reasons. First, they seem least likely to violate the independence assumption. The observation that children in private schools tend to come from higher-income families would seem to imply we should expect a negative correlation between income and strength of preferences for local public school quality.[29] This is less likely to be true for retired households who have fewer school-age children. There is also no obvious reason to expect correlation between their income and preferences for other public goods. Poor air quality should affect retirees' health regardless of income. The second strategic advantage of using retired households is that they bridge the single and dual-market versions of the model. Generalizing the urban landscape to include labor markets does not affect the choice set faced by retirees; their income is fixed. Since retirees choose from the same 122 communities in both versions of the model, both models should return the same information about their preferences. This requires both models to produce the same estimates for ρ and ξ, which is

[29] Within the study region, the average income of households with children enrolled in private schools is 42% higher than for those enrolled in public schools according to year 2000 Census School District data.

guaranteed if they are estimated from data on retired households.

To implement the second-stage of the estimation, all households were classified according to 10 income bins reported in the Census data, and each household was assigned a level of income equal to the midpoint of its bin.[30] Then, the objective function used to estimate ρ was defined as the sum of the difference in the marginal distributions of $F(\alpha, \gamma_{air}, \gamma_{school}, \gamma_{\xi})$ for all pairwise combinations of income for retired households. The function was minimized using a grid search over [-.6, 0], which includes the range of estimates from previous studies. The function was minimized at $\rho = -0.131$. While this estimate is 10 times as large as the result from the vertical model (-0.013), they imply similar values for the elasticity of substitution between public and private goods. Here, the elasticity is 0.88 compared to 0.99 in the vertical case.[31]

Estimates for the distribution of unobserved public goods are also very similar between the vertical and horizontal models. The average community differs by 3 places in the ranking by ξ between the two models. Overall, ξ becomes increasingly important in explaining location choices as one moves closer to the San Francisco Bay. Some of the unobserved public goods that seem likely to be influencing the spatial pattern of ξ include climate, open space, and cultural amenities. The San Francisco Bay Area generally has the mildest weather in the study region and the most opportunities for dining and nightlife. The Bay Area also has a relatively large share of land in open space. The San Francisco, San Jose, and Santa Cruz PMSAs have the highest median values for ξ and the largest share of land in state parks. This pattern is consistent with previous sorting applications which have found open space to be an important determinant of where households locate (Walsh [2007]).

Using the estimates for ρ and ξ, the Gibbs algorithm recovered an approximation to the partition of preference space defined by 1,220,000 points—1000 points drawn from each of the 122 regions at 10 different levels of income.[32] Recall from section 2

[30] Measured in thousands, the midpoints are: [5 12.5 22.5 35 45 55 67.5 87.5 112.5 175].
[31] The elasticity of substitution is defined as: $\sigma = 1/(1-\rho)$.
[32] This followed a burn-in of 100 draws to reduce sensitivity to starting values.

that the logic of revealed preferences may not fully bound regions that correspond to locations with extreme provision of public goods. Therefore, absolute upper and lower bounds had to be imposed on each dimension to ensure that the points were drawn from the "economically relevant" portion of the unbounded regions.

The job skill parameter (θ) was bounded by 0 and 1.5. Its lower bound implies the worker's idiosyncratic skills prevent them from gaining employment in any location other than their current niche, whereas its upper bound implies the worker is overqualified at their current job and could make 150% of the market wage in alternative job locations. The lower bound on ϕ was set to 0, allowing a worker's implicit opportunity cost of time spent commuting to be 0, and the upper bound was set so that the worker's opportunity cost of time equals their wage rate.[33] The weights in the public goods index were normalized to sum to 1, allowing the bounds for α to be set based on prior assumptions about the range of plausible values for the MWTP. The lower bound on α was set to 0, restricting MWTP for public goods to be nonnegative. Its upper bound was set to correspond to a \$500 MWTP for improved air quality. More precisely, the upper bound on α sets a \$500 limit on an individual household's willingness-to-pay for a 1 part per billion (ppb) reduction in the annual average of the top 30 1-hour daily maximum readings for ozone concentrations. This measure is not directly comparable with estimates for the MWTP in much of the existing literature where air quality is typically measured by particulate matter or by the number of days during a year that ozone levels exceed state or federal standards. However, to the extent that all of these measures are simply different proxies for clean air, they can be compared in terms of a common proportionate change. Sieg et al. (2004) use this logic to translate the range of estimates for the average MWTP in the existing literature into measures that would be comparable to the willingness-to-pay for a 1.5 ppb reduction in ozone concentrations. Converted to year 2000 dollars, the range is \$11 to \$231. Measured in these normalized units, the upper bound on α would

[33] For example, at the upper bound a worker who is paid \$20/hour would be exactly indifferent between their current location and another location that would increase their commute by 1 hour and raise their wage by \$20.

imply a value of $750. Thus, the upper bound limits an *individual* household's MWTP to roughly triple the upper bound on *average* MWTP in the existing literature.

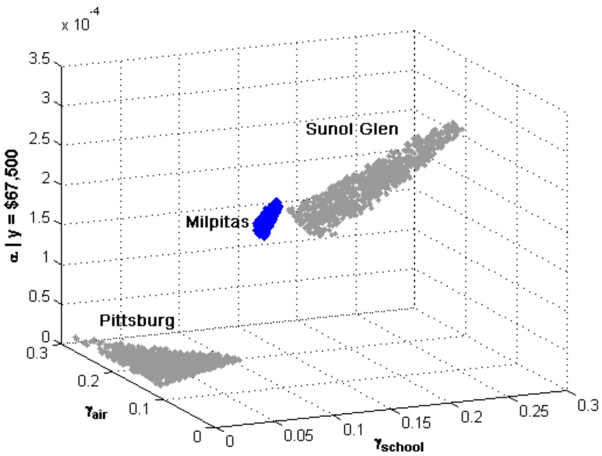
The resulting partition generalizes the revealed preference logic from the vertical model. This can be seen by comparing the preference regions that each model assigns to households living in three communities—Pittsburg, Milpitas, and Sunol Glen. Of the three, Sunol Glen and Milpitas provide more of every public good than Pittsburg.[34] Therefore, regardless of relative preferences, every household will perceive Pittsburg as providing the lowest quality bundle of public goods. Given this unanimous ordering, a household's choice to live in Pittsburg reveals that they have weaker preferences for public goods relative to private goods compared to households with the same income in the other two communities. This logic is reflected by the stratification of households in figure 4 and figure 5A. In both figures, the preference sets for Sunol Glen and Milpitas lie above the set for Pittsburg in the α dimension. However notice that, unlike figure 4, households in Sunol Glen and Milpitas have overlapping ranges of values for α in figure 5. This occurs because the two communities are not strictly ordered by their provision of public goods. Sunol Glen has higher quality schools and Milpitas has cleaner air. Otherwise they are nearly identical; the price of housing and provision of ξ differ by approximately 1% between the two communities. Thus, the choice between Sunol Glen and Milpitas helps to identify households' preferences for air quality relative to school quality. This logic underlies the result in figure 5B that households in Sunol Glen have strictly higher relative preferences for school quality.

More generally, the size and shape of each preference region reflects the substitution possibilities available to the households in the corresponding community.[35] Preferences are better identified for households that live in communities with closer substitutes. For example, there are at least five other communities that are very similar to Milpitas in

[34] The normalized values for {air quality, school quality, ξ, price} in each community are as follows: Pittsburg {0.82, 0.79, 0.17, 1.42}; Milpitas {0.96, 1.05, 0.54, 2.61}; Sunol Glen {0.91, 1.20, 0.53, 2.62}.

[35] While the Gibbs algorithm sampled uniformly over each preference region, there appears to be sparseness near some of the edges in figure 5B. For example, see the upper left corner of Pittsburg. This is because its preference region is approximately pyramidal and the sparseness occurs in the tip which would be consistent with a uniform density of points.

their provision of air and school quality. Consequently, Milpitas has a small preference region compared to Pittsburg and Sunol Glen which have fewer close neighbors in public goods space.



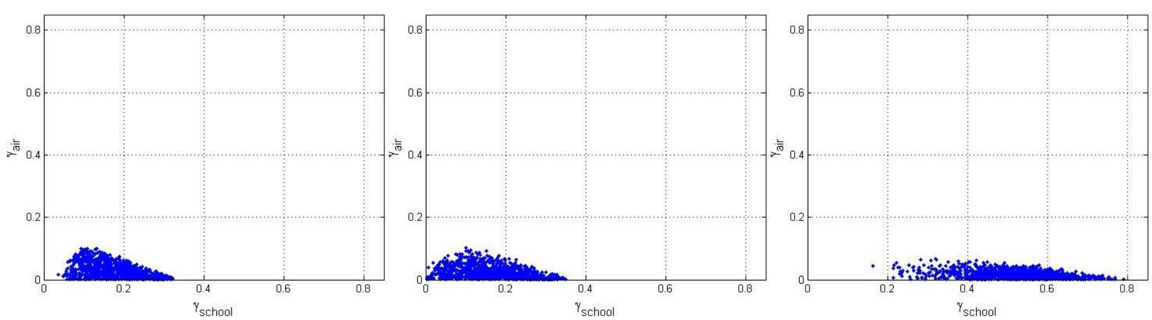
A. α, γ *space* | *y=$67,500* B. *Projection in* γ *space*

FIGURE 5.—Preference Regions for 3 Communities, Single-Market Horizontal Case

5.4. *Second Stage Estimation Results: Dual-Market Horizontal Model*

In the dual-market version of the model, the approximation to the partition of preference space is defined by 58,960,000 points—1000 points drawn from each of the 268 regions for each of the 220 (occupation, non-wage income) pairs. The main difference from the single-market partition is that adding work destinations to the choice set expands the borders of the preference sets. Intuitively, heterogeneity in job skill and the opportunity cost of time provide new ways to explain observed location choices.

Figure 6 provides a representative example of how the preference regions differ. Panels A, B, and C project the preference sets recovered for architects and engineers in the Acalanes school district onto $\gamma_{air}, \gamma_{school}$ space. In the single-market case (panel A) the choice to live in Acalanes reveals strong preferences for school quality relative to air quality because Acalanes has high quality schools (90th percentile) and low quality air (14th percentile). Of all the possible job destinations for architects and engineers who live there, the Oakland PMSA requires the shortest commute (24 minutes). Therefore, the choice to live in Oakland may reveal a high opportunity cost of time rather than strong preferences for school quality. This possibility is reflected in the way the preference region in panel B is "stretched" to the left compared to panel A. The lowest values for γ_{school} correspond to high values for the opportunity cost of time parameter (ϕ). In contrast, the preference region is stretched to the right for workers who make the relatively long commute to San Francisco (55 minutes). In this case, the highest values for γ_{school} are paired with low values for the job mobility parameter (θ). For an architect or engineer who is "stuck" working in San Francisco, the choice to live far from their job reveals strong preferences for the public goods provided by that community—in this case school quality.



A. *Acalanes* (*single-market*)

B. *Acalanes* → *Oakland* (*24 mins, wage=$57,700*)

C. *Acalanes* → *San Francisco* (*55 mins, wage=$58,500*)

FIGURE 6.—Stratification by Relative Preferences with and without Job Choices

To make a more general comparison between the single and dual-market partitions, each was translated into a distribution of preferences by sampling uniformly over each region according to the population of households in the corresponding community.[36] For example, the census data report 232 households with a primary earner in the architecture and engineering occupation who live in the Acalanes school district, work in the Oakland PMSA, and have total income of $112,500. Therefore, 232 draws were chosen uniformly from the region of the partition that corresponds to this household "type".

[36] This does not imply preferences are uniformly distributed within the *population* of households.

This process was repeated for every household type so that the resulting distributions represent all 3.2 million households in the study region. Table IV reports means and standard deviations that describe the marginal distribution of each parameter.

In the dual-market case, the means for α, γ_{air} and γ_{school} are all slightly larger and the mean for γ_{ξ} is slightly smaller. Intuitively, without job opportunities to help explain location choices, the single-market version of the model has to assign more importance to unobserved public goods to rationalize observed behavior. The larger standard deviations on γ_{air}, γ_{school}, and γ_{ξ} in the dual-market case reflect the way that job opportunities tend to widen the bounds on the preference regions. The mean value for θ suggests a high degree of geographic job specialization; it implies the average worker would earn approximately half of the market wage if they were to change job locations. Another interpretation would be that this relatively low value reflects a high job search cost. Finally, using ϕ to calculate the opportunity cost of time spent commuting suggests that the average worker values leisure time at 39% of their wage rate. This is quite similar to the rule-of-thumb (33%) that is often used in recreation demand studies (Phaneuf and Smith [2005]).

TABLE IV
Mean (standard deviation) for Distributions of the Heterogeneous Parameters

Model	Distributional Assumption	Parameter					
		log (α)	γ_{school}	γ_{air}	γ_{ξ}	θ	φ
Single-Market Vertical	f(α,y)~lognormal	1.132 (0.539)	1.000	0.149	----	----	----
Single-Market Horizontal	f(α,γ|y)~uniform in each preference set	-8.184 (4.309)	0.114 (0.153)	0.143 (0.142)	0.744 (0.206)	----	----
Dual-Market Horizontal	f(α,γ,φ,θ|y)~uniform in each preference set	-8.079 (4.286)	0.159 (0.196)	0.167 (0.180)	0.674 (0.268)	0.461 (0.275)	1.478 (1.258)

Table IV also reports the point estimates for α and γ_{air} from the vertical model. They are not comparable to the horizontal results in terms of magnitude since they correspond to different estimates for ρ and ξ. Nevertheless, there is a striking difference between the relative values for the (average) weights estimated for the horizontal model and the (constant) weights estimated for the vertical model. The ratio of γ_{air} to γ_{school} in the two horizontal models is an order of magnitude larger than in the vertical case. This could be due to the many differences between the two estimators, or it could simply reflect the large standard error on the vertical point estimate for γ_{air}.

In summary, the results from each of the three sorting models can be used to characterize the distribution of preferences for public goods in the population of households who live in the San Francisco-Sacramento area. The three models differ in how they define a locational equilibrium, how they depict heterogeneity in households, and in the restrictions they place on the shape of the distributions used to characterize sources of heterogeneity. The differences in these identifying assumptions lead to substantial differences in the information recovered about preferences, as illustrated by the summary statistics in table IV and the shape of the partitions in figures 4, 5, and 6.

5.5. *Implications: Marginal Willingness-to-Pay for Improved Air Quality*

To compare the economic implications of the three models, the information about preferences was translated into distributions of the willingness-to-pay for a marginal (1 ppb) reduction in ozone concentrations. For the vertical model, this simply requires drawing a sample of households from the joint distribution of income and preferences defined by the parameter estimates for $\mu^{\alpha}, \mu^{y}, \sigma^{\alpha}, \sigma^{y}, \lambda$ and converting each draw into the corresponding MWTP. Likewise, the horizontal partitions were translated into distributions of MWTP by sampling from each region of preference space according to the associated population of households and then converting each draw into the corresponding MWTP. This approach was used to generate three distributions. First, the assumption that preferences are distributed uniformly *within* each preference region was translated into a distribution of MWTP. Then, upper and lower bounds on that distribution were generated. For example, the lower (upper) bound distribution was constructed by assigning every household the lowest (highest) possible MWTP that would be consistent with its observed location choice. Any assumption about the shape of the joint distribution of preferences will lead to a distribution of MWTP that falls within these bounds.

The difference between the upper and lower bound distributions can be used to measure the economic significance of assumptions on the distribution of preferences. Table V reports the share of households within 7 different "identification intervals". For example, the difference between the highest and lowest MWTP that would be consistent with observed location choices lies between \$0 and \$10 for 3.1% of households in the single-market case. In other words, the MWTP is identified to within \$10 for these households. Likewise, the MWTP is identified to within \$25 for 16.3% of households (3.1% + 13.2%). Moving from the single to the dual-market case decreases the share of households for whom the MWTP is precisely estimated. This is consistent with the observation that the dual-market preference regions typically have wider bounds.[37]

TABLE V
Identifying MWTP for Improved Air Quality, Horizontal Models

Model	Share of Households with | max (MWTP) - min (MWTP) | in the Range:						
	\$0-\$10	\$10-\$25	\$25-\$50	\$50-\$75	\$75-\$100	\$100-\$250	\$250-\$500
Single-Market	3.1%	13.2%	22.0%	17.9%	11.6%	20.2%	12.1%
Dual-Market	2.5%	7.2%	12.6%	12.3%	9.1%	22.9%	33.3%

Table VI provides summary measures of the MWTP distributions and compares them to the corresponding results from the vertical model. The range of estimates for average per/household MWTP in the dual market case (\$41 to \$237) contains the range in the single-market case (\$67 to \$180) which contains the point estimate from the vertical model (\$113). This illustrates a type of "bias/variance" tradeoff. If the depiction of utility in the dual-market case represents the "truth", then treating income as exogenous and preferences as vertically differentiated will have two effects. It will bias the resulting

[37] The upper bound of \$500 that was imposed on the two horizontal models truncates the preference regions for approximately 8.1% of households in the single-market case, compared to 22.1% in the dual-market case.

welfare measures and it will decrease the sensitivity of those measures to assumptions on the distribution of heterogeneous preference parameters. The table also illustrates another general feature of the results: conditional on the uniform assumption, introducing horizontal differentiation and accounting for job opportunities both tend to increase the MWTP.

TABLE VI
Average per/household MWTP for Improved Air Quality, 3 Sorting Models

Single-market, Vertical	Single-market, Horizontal			Dual-market, Horizontal		
	min MWTP	uniform pref.	max MWTP	min MWTP	uniform pref.	max MWTP
113	67	125	180	41	142	237

Compared to the results from reduced-form hedonic studies of the housing market, the dual-market estimates for the MWTP are relatively high. Converting the range of normalized values for the existing literature into measures that would be equivalent to the average MWTP for a 1 ppb ozone reduction implies a range from $7 to $154 (year 2000 dollars). The higher range produced by the dual-market estimator ($41 to $237) could stem from methodological differences or simply from differences in the study region. The $7 and $154 estimates are both for Los Angeles which has much higher ozone concentrations than the San Francisco-Sacramento area (Seig et al. [2004]. Moreover, median income in the San Francisco CMSA is 35% higher than in the Los Angeles CMSA. If Northern and Southern California were considered as part of the same choice set, the relationship between MWTP, air quality, and income would imply that households in San Francisco and Sacramento would tend to have a higher MWTP than those in Los Angeles.

The result that MWTP tends to increase when the choice set is expanded to include job opportunities is consistent with the interregional hedonic literature which has found that housing prices and wage rates both reflect a substantial share of the implicit price of environmental amenities (e.g. Roback [1982], Blomquist et al. [1988]). These applications estimate quality-of-life indices under the assumptions that households have homogeneous preferences for public goods and are freely mobile in national markets for housing and labor. Bayer, Keohane and Timmins (2006) extend this literature to relax the free mobility assumption and estimate the MWTP for air quality. Their analysis includes a cost for moving between states but, like the present study, they treat households as freely mobile within each state. Converting the point estimate from their national model to a measure that would be comparable to a 1 ppb ozone reduction in the San Francisco-Sacramento region would imply a MWTP of $224 (year 2000 dollars).[38] This figure falls within the range of results reported in table VI for the average MWTP in the dual-market case despite numerous methodological differences between the two studies. For example, Bayer et al. use metropolitan statistical areas as their spatial unit of observation and restrict households to have identical preferences for air quality.

Finally, from a methodological perspective, the closest comparison to the existing literature is to Sieg et al's (2004) application of the single-market vertical model to Los Angeles in 1990. They report an average MWTP of $66. However, the average level of ozone concentrations across the communities in their application is 150 ppb, compared to a maximum of 109 here. The Sacramento PMSA provides the closest approximation to the income and ozone conditions in Los Angeles. The average level of ozone concentrations for the communities physically located in Sacramento is 94 ppb and the median income is 1.5% higher than in Los Angeles. For the households who live in these communities, the average MWTP predicted by the single-market vertical model is $23, compared to $73 and $82 for the two horizontally differentiated models (under the uniform assumption). The low estimate for the vertical model reflects the fact that the communities in the Sacramento PMSA have the lowest housing prices in the study region. Therefore, conditional on income, they are assigned the lowest values for α, which imply the lowest values for the MWTP. The horizontal models also assign relatively low values to households in these communities, but recognize that variation in relative preferences and job opportunities may induce some households with relatively strong preferences for air

[38] The $224 figure is calculated using the WTP elasticity for particulate matter from the full version of their model (0.34). This calculation assumes particulate matter and ozone can be compared in terms of a common proportionate change.

quality to locate there.

6. SUMMARY

Tiebout (1956) suggested that consumers reveal their preferences for local public goods by the residential locations they choose. While this logic forms the basis for empirical models of sorting behavior, recent evidence suggests that local public goods are not the dominant factor in residential choice. Rhode and Strumpf (2003) find that as moving costs declined between 1850 and 1990, U.S. counties and municipalities became *less* stratified by public goods provision and household demographics—the opposite of what we would expect from a model of pure Tiebout sorting. Likewise, the American Housing Survey consistently reports "convenient to job" as the reason most frequently cited by households for choosing to live in their current neighborhood.[39] If residential choice is driven by job opportunities, ignoring the possibilities for labor market adjustment may systematically bias estimates of household preferences for local public goods.

This paper has developed a new structural estimator of household preferences for local public goods based on a model of impure Tiebout sorting in which job opportunities limit residential mobility. By redefining each location as a job-house combination and recognizing job skill as an additional dimension of heterogeneity, the model has extended the existing sorting literature to consider a dual-market locational equilibrium. This framework recognizes that each working household faces a limited set of job options. They may be forced to choose between lower-amenity communities with cheaper housing and better access to high-paying jobs and communities with higher amenities, poorer access, and more expensive housing. The choices made by households facing this tradeoff reveal features of their preferences.

The dual-market framework generalizes two features of Epple and Sieg's (1999) estimable model of Tiebout sorting. First, wage income and leisure time are both endogenous to location choice. This makes it possible to extend Tiebout's logic to consider

[39] For examples see table 2-11 of the surveys for 1997, 1999, 2001, 2003, 2005, and 2007. In each year, "convenient to job" is the most frequently cited reason for "choice of present neighborhood" as well as the most frequently cited specific choice for "main reason for choice of present neighborhood".

sorting behavior in differentiated labor markets. Ceteris paribus, heterogeneous workers are assumed to select a job based on the wages they can earn and the required commute time. The model also relaxes vertical differentiation to allow households to differ in their relative preferences for leisure time and multiple public goods. Thus, in a locational equilibrium, working households are simultaneously sorted among the housing and labor markets according to their heterogeneous preferences and skills. Opportunities for adjustment in both markets make the implicit cost of consuming local public goods a function of housing prices, wage rates, and commute times.

Modeling the interrelated choices made by heterogeneous households in multiple markets presents new challenges for the structural estimation of preferences. House locations, job locations, and commuting patterns must all be observed, and the wage rate a household would earn at each job location must be specified. There is also greater scope for structural assumptions to influence results. In particular, expanding the dimensionality of unobserved heterogeneity increases the sensitivity of welfare measures to assumptions about the shape of the distributions used to characterize sources of heterogeneity. A key feature of the new estimator is that it avoids the need to impose prior assumptions on the joint distribution of random parameters. Instead, preference space is partitioned into mutually exclusive cells that explain observed behavior. This partition can be used to recover bounds on the distribution of preferences and bounds on the range of consistent welfare measures for a given policy change.

The impact of moving from Epple and Sieg's model to the new dual-market estimator, in terms of the average MWTP for air quality in Northern California, ranges from a 64% decrease to a 110% increase, depending on distributional assumptions. Under the assumption that preferences are uniformly distributed within each cell of the partition, moving from vertical to horizontal differentiation increases estimates for average MWTP. Accounting for job opportunities leads to a similar increase. The increase in average MWTP under the uniform assumption together with the increased sensitivity of that result to extreme distributional assumptions illustrates a type of bias/variance tradeoff that applies generally to microeconometric models of choice among differentiated alternatives.

In the context of Tiebout's (1956) revealed preference logic, this tradeoff implies that while households signal their preferences for local public goods by the residential (and job) locations they choose, what we infer from their choices depends on our understanding of the ways in which people differ.

REFERENCES

Anderson, S. P., A. de Palma, and J. Thisse (1992): *Discrete Choice Theory of Product Differentiation,* Cambridge, Massachusetts: The MIT Press.

Bajari, P. and C. L. Benkard (2005): "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach," *Journal of Political Economy,* 113, 1239-1276.

Bayer, P., N. Keohane, and C. Timmins (2006): "Migration and Hedonic Valuation: The Case of Air Quality," NBER Working Paper #12106.

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

Bayer, P., R. McMillan, and K. Reuben (2005): "An Equilibrium Model of Sorting in an Urban Housing Market," NBER Working Paper #10865.

Bernheim, B. D., and A. Rangel (2009): "Beyond Revealed Preference: Choice Theoretic Foundations for Behavioral Welfare Economics," *Quarterly Journal of Economics*, 124, forthcoming.

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

Berry, S., and A. Pakes (2007): "The Pure Characteristics Demand Model," *International Economic Review*, 48, 1193-1225.

Blomquist, G. C., M. C. Berger, and J. P. Hoehn (1988): "New Estimates of Quality of Life in Urban Areas," *American Economic Review*, 78, 89-107.

Briesch, R. A., P. K. Chintagunta, and R. L. Matzkin (2007): "Nonparametric Discrete Choice Models with Unobserved Heterogeneity," *Mimeo*, Northwestern University.

Calabrese, S., D. Epple, T. Romer, and H. Sieg (2006): "Local Public Good Provision: Voting, Peer Effects, and Mobility," *Journal of Public Economics*, 90, 959-981.

Chernozhukov, V., and C. Hansen (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73, 245-261.

Chernozhukov, V., G. W. Imbens, and W. K. Newey (2007): "Instrumental Variable Estimation of Nonseparable Models," *Journal of Econometrics*, 139, 4-14.

Chesher, A. (2003): "Identification in Nonseparable Models," *Econometrica*, 71, 1405-1441.

Epple, D. and H. Sieg (1999): “Estimating Equilibrium Models of Local Jurisdiction,” *Journal of Political Economy*, 107, 645-681.

Epple, D., M. Peress, and H. Sieg (2005): “Identification and Semiparametric Estimation of Equilibrium Models of Local Jurisdiction,” Mimeo, Carnegie Mellon University.

Epple, D. and G. J. Platt (1998): “Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Incomes,” *Journal of Urban Economics*, 43, 23-51.

Feenstra, R. C., and J. A. Levinsohn (1995): “Estimating Markups and Market Conduct with Multidimensional Product Attributes,” *Review of Economic Studies*, 62, 19-52.

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

Geweke, J. F. (1996): “Monte Carlo Simulation and Numerical Integration,” in *Handbook of Computational Economics*, ed. by H. Amman, D. Kendrick and J. Rust, Amsterdam: North-Holland.

Imbens, G. W., and W. K. Newey (2008): "Identification and Estimation of Triangular Simultaneous Equation Models without Additivity," Mimeo, MIT.

Matzkin, R. L. (2003): “Nonparametric Estimation of Nonadditive Random Functions,” *Econometrica*, 71, 1339-1375.

O’Sullivan, A, T.A. Sexton, and S.M. Sheffrin (1995): *Property Taxes and Tax Revolts: The Legacy of Proposition 13*. Cambridge: Cambridge University Press.

Palmquist, R. B. (2005): “Weak Complementarity, Path Independence, and the Intuition of the Willig Condition,” *Journal of Environmental Economics and Management*, 49, 103-115.

Phaneuf, D. J. and V. K. Smith (2005): “Recreation Demand Models,” in *Handbook of Environmental Economics,* Vol. II, ed. by K. Göran-Mäler and J. Vincent, Amsterdam: North Holland.

Polinsky, M. A. (1977): "The Demand for Housing: A Study in Specification and Grouping," *Econometrica*, 45, 447-462.

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

Rhode, P. W. and K. S. Strumpf. (2003): "Assessing the Importance of Tiebout Sorting: Local Heterogeneity from 1850 to 1990," *American Economic Review*, 93, 1648-1677.

Roback, J. (1982): "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 90, 1257-1278.

Rosen, S. (1979): "Wage-Based Indexes of Urban Quality of Life," in *Current Issues in Urban Economics*, ed. by P. Mieszkowski and M. Straszheim, Baltimore: John Hopkins University Press.

Scotchmer, S. (1985): "Hedonic Prices and Cost/Benefit Analysis," *Journal of Economic Theory*, 37, 55-75.

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

______(2002): "Interjurisdictional Housing Prices in Location Equilibrium," *Journal of Urban Economics*, 52, 131-153.

Smith, V. K. and H. S. Banzhaf (2004): "A Diagrammatic Exposition of Weak Complementarity and the Willing Condition," *American Journal of Agricultural Economics*, 86, 455-466.

Starrett, D. A. (1981): "Land Value Capitalization in Local Public Finance," *Journal of Political Economy*, 89, 306-327.

Tiebout, C. M. (1956): "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 64, 416-424.

Walsh, R. L. (2007): "Endogenous Open Space Amenities in a Locational Equilibrium," *Journal of Urban Economics*, 61, 319-344.