

**THE NEW ECONOMICS OF EQUILIBRIUM SORTING
AND POLICY EVALUATION USING HOUSING MARKETS**

Nicolai V. Kuminoff

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

V. Kerry Smith

Regents' Professor and
W. P. Carey Professor
Department of Economics
Arizona State University
kerry.smith@asu.edu

Christopher Timmins

Professor
Department of Economics
Duke University
timmins@econ.duke.edu

December 2012

THE NEW ECONOMICS OF EQUILIBRIUM SORTING AND POLICY EVALUATION USING HOUSING MARKETS

ABSTRACT: Households “sort” across neighborhoods according to their wealth and their preferences for public goods, social characteristics, and commuting opportunities. The aggregation of these individual choices in markets and in other institutions influences the supply of amenities and local public goods. Pollution, congestion, and the quality of public education are examples. Over the past decade, advances in economic models of this sorting process have led to a new framework that promises to alter the ways we conceptualize the policy evaluation process in the future. These “equilibrium sorting” models use the properties of market equilibria, together with information on household behavior, to infer structural parameters that characterize preference heterogeneity. The results can be used to develop theoretically consistent predictions for the welfare implications of future policy changes. Analysis is not confined to marginal effects or a partial equilibrium setting. Nor is it limited to prices and quantities. Sorting models can integrate descriptions of how non-market goods are generated, estimate how they affect decision making and, in turn, predict how they will be affected by future policies targeting prices or quantities. Conversely, sorting models can predict how equilibrium prices and quantities will be affected by policies that target product quality, information, or amenities generated by the sorting process. These capabilities are just beginning to be understood and used in applied research. This survey article aims to synthesize the state of knowledge on equilibrium sorting, the new possibilities for policy analysis, and the conceptual and empirical challenges that define the frontiers of the literature.

1. *Introduction*

Economists use sorting as a metaphor for the way that market forces partition economic agents across segments of a market. Households “sort” across neighborhoods according to their wealth and their preferences for housing characteristics, local public goods, social characteristics of their neighbors, and commuting opportunities. Workers “sort” across jobs according to their qualifications and preferences for job attributes. In situations with other differentiated products such as automobiles, breakfast cereal, and computers, we expect that consumers who have similar preferences and face similar constraints will make similar choices. This sorting process reveals information about consumers, and firms have learned to exploit it to increase their profits. They design differentiated products and set prices to take advantage of what is known about consumer heterogeneity. Knowledge of consumer heterogeneity can also be used to evaluate past policies and design new ones. This is especially true for policies targeting public goods and externalities. It is important to distinguish sources of preference heterogeneity that can be related to observable features of consumers, their constraints, and their choices from that which is not observable by an analyst. The challenge for economists is to integrate all the sources of heterogeneity into a description of sorting behavior. The models need to reflect the information available to the agents involved and their constraints. Ideally, they would describe how agents respond to exogenous events and what the collection of those actions implies for market and non-market outcomes.

“Equilibrium sorting” models use the properties of market equilibria, together with information on the behavior of economic agents, to estimate structural parameters that characterize agent heterogeneity. The results can be used to develop theoretically consistent measures of the welfare implications of future policy changes. Analysis is not confined to marginal effects or a partial equilibrium setting. Nor is it limited to prices and quantities. As heterogeneous agents sort, their collective behavior can influence the supply of amenities. These adjustments can be represented as part of the characterization of the equilibria. Pollution, congestion, and opportunities for social interaction provide examples where an endogenous change in a non-market attribute can be derived as part of the equilibrium adjustment. Sorting models can integrate descriptions of how these amenities are generated, estimate how they affect decision making and, in turn, predict how they will be affected by future policies targeting prices or quantities. Conversely, sorting models can predict how equilibrium prices and quantities will be affected by

policies that target product quality, information, or amenities generated by the sorting process. These capabilities are just beginning to be understood and used in applied research.

Sorting models build on the intellectual foundations of the literature on hedonic and discrete-choice models of differentiated product markets. They combine the information provided by an *equilibrium* hedonic price function (Sherwin Rosen 1974; Dennis Epple 1987; Ivar Ekeland, James J. Heckman, and Lars Nesheim 2004) with a formal description for the choice process that underlies market *sorting* of heterogeneous agents (Daniel McFadden 1974; Timothy Bresnahan 1987; Steven Berry, James Levinsohn, Ariel Pakes 1995). This *equilibrium sorting* framework can depict a mixture of discrete and continuous choices made by a population of heterogeneous agents, while recognizing that characteristics of the objects of choice may be determined endogenously (Epple and Holger Sieg 1999; Patrick Bayer and Christopher Timmins 2005, 2007).

What ideas distinguish the economics of equilibrium sorting from past strategies for modeling differentiated goods? First, in addition to characterizing sources of unobserved heterogeneity such as technology and preferences, they include a wide array of observable features that distinguish economic agents. These observable dimensions of agent heterogeneity can be used in descriptions of sorting behavior and are often important in characterizing the implications of policies. Through the sorting process, that heterogeneity affects the endogenously determined attributes of the choice alternatives available to agents. In the housing market, for example, the attributes of the neighborhoods that a household chooses from may depend on where its primary earner works, while its preferences for school districts may depend upon the levels of education attained by the adult members of the household. As households with different incomes and levels of education decide where to live, they will influence the demographic compositions of neighborhoods. When households vote, their preferences will shape public policies that influence local public goods such as school quality or protected open space. Other amenities may be determined endogenously as an indirect outcome of the sorting process. This allows for the evaluation of the distributional implications of a simulated equilibrium.

Endogenous amenities create challenges for estimating these models as well as affect market outcomes for private goods, such as housing prices and wage rates. This creates a second distinction between the economics of equilibrium sorting and earlier models of the demand for

differentiated products. The sorting literature seeks to incorporate the “general equilibrium” feedback effects between economic agents and their environments. A shock to the housing market that induces a change in residential location patterns may lead to a redistribution of local amenities that induces more migration and housing development, which continues until prices adjust and markets clear. Modeling these feedback effects is important for researchers interested in simulating the impacts of a counterfactual policy.

Third, the equilibrium sorting literature considers how public policies can be designed to exploit what we learn about forms of heterogeneity, endogeneity, and feedback. Some time ago, Alan S. Blinder and Harvey S. Rosen (1985) demonstrated how information about preference heterogeneity could, in principle, be used to design more efficient taxes on private goods. Emmanuel Saez (2010) recently identified a testable implication of “bunching” at kinks, based on workers’ responses to nonlinearities in the tax code. Equilibrium sorting models provide the means to implement both the original Blinder-Rosen idea and the Saez test and extend them for policies that target public goods or other amenities that affect agents differently.

Agent heterogeneity, endogenous attributes, and feedback effects can all have first-order policy implications (Sieg et al. 2004; V. Kerry Smith et al. 2004; Maria Marta Ferriera 2007; Timmins 2007; Randall L. Walsh 2007; H. Spencer Banzhaf and Walsh 2008; Nicolai V. Kuminoff 2009; H. Allen Klaiber and Daniel J. Phaneuf 2010b; Constant I. Tra 2010). Market equilibria can depend on feedback effects that occur through *non-market* transmission routes. For example, in Walsh (2007) household utility is enhanced by living in neighborhoods with open space. However, the amount of open space declines as more land is allocated to houses in each neighborhood. This feedback effect implies that each household’s location choice depends on the choices made by other households. The demand side of a sorting equilibrium that clears this market is itself a Nash equilibrium that fits within the class of aggregative public goods games characterized by Richard Cornes and Roger Hartley (2007). Sorting in response to feedback leads to a surprising result in Walsh’s policy simulation. Increasing the amount of land in public preserves can actually decrease the total amount of land in open space in the metro area. The mechanisms that produce this outcome mirror a counterintuitive result from Matthew J. Kotchen’s (2006) theoretical model of Nash equilibria in green markets.

Housing markets provided a natural testing ground for equilibrium sorting models. The

techniques developed in this context have been used to study behavior in a wide range of differentiated product markets. Recent applications have considered network effects (Marc Rysman 2004; Shanjun Li 2012), location choices of firms (Katja Seim 2006), markets for education (Epple, Richard Romano, and Sieg 2006, 2012), social interactions in labor markets (Bayer, Steve Ross, and Giorgio Topa 2008), and the impact of congestion on recreation demand (Timmins and Jennifer Murdock 2007).

The potential for using equilibrium sorting models to conduct high-resolution policy analysis is intriguing, but are their assessments reliable? The barriers to recovering structural parameters are formidable. Identifying parameters describing household preferences first requires collecting micro data that accurately describe the choice alternatives and financial constraints that enter a household's decision process. In addition to the standard concerns about misspecification and omitted variable bias, the general equilibrium structure of a sorting model may suggest the need to develop instruments for endogenous amenities, such as open space or school quality. Additional challenges arise when sorting models are used to predict general equilibrium responses to counterfactual policies. If a sorting model has multiple equilibria, then the analyst should determine whether the model's policy-relevant predictions are robust to the choice among equilibria. Moving costs and other important sources of friction should be included among the constraints; forward-looking behavior and the role for dynamics should be considered; and so on. Considerable progress has been made toward addressing some of these challenges, while others continue to provide opportunities for research. Addressing these challenges is a worthy goal for empirical work.

Some of the recent approaches have come from adapting insights from the econometric literature on program evaluation (Guido M. Imbens and Jeffrey M. Wooldridge 2009). For example, Bayer, Fernando Ferreira and Robert McMillan (2007) adapt the logic of regression discontinuity design to mitigate omitted variable bias in their estimates for structural preference parameters. Other studies have illustrated how the internal validity of a sorting model can be improved by using theory and quasi-experimental sources of variation in the data to guide the development of instrumental variables (Epple and Sieg 1999, Bayer and Timmins 2007, Bayer, Keohane, and Timmins 2009). Recent work has also begun to tackle the difficult, but extremely important, task of assessing external validity. A recent study, by Sebastian Galiani, Alvin

Murphy, and Juan Pantano (2012), uses a sorting model to make out-of-sample predictions about participation in a housing assistance program targeting low-income households. Remarkably, their predictions align very closely with actual participation observed in a randomized controlled trial of the “Moving to Opportunity” experiments conducted by the U.S. Department of Housing and Urban Development. This study is an example of research that will help to increase confidence in the circumstances when sorting models can reliably inform future policy choices. The early evidence is encouraging, but far from conclusive. More work is needed.

This survey aims to synthesize the state of knowledge on equilibrium sorting, the new possibilities for policy analysis, and the conceptual and empirical challenges that define the frontiers of the literature. We concentrate on research in public and environmental economics, with particular attention to the market for housing. Models of household location choice build on theory and methods developed in related fields, especially industrial organization and labor economics. We highlight connections to work in those areas without providing a comprehensive assessment. Our focus is on recent research. While we do not present a complete historical perspective, much of the work we cover was influenced by the seminal papers written by Charles M. Tiebout (1956), William Alonso (1964), Edwin S. Mills (1967), Richard F. Muth (1969), Wallace E. Oates (1969), Thomas C. Schelling (1969), and Edwin T Haefele (1971).

Our survey begins by describing the foundations of the new equilibrium sorting literature in section 2, from early median voter models of the demand for public goods to the modern discrete choice framework for describing how households sort over neighborhoods. Section 3 covers the evolution of sorting theory. This line of research has sought to characterize multi-community equilibria with peer effects, voting, and other forms of social interaction. The implied relationship between property values, housing characteristics, and local public goods can be described by a hedonic price function. Empirical models use properties of sorting equilibria to recover household preferences and estimate the demand for public goods. Section 4 covers the leading empirical models of sorting behavior. In section 5 we discuss how these models change the way we evaluate public policy. Leading examples are provided (education, air pollution, and land use) and we comment on how the new sorting models can improve future evaluations. Finally, section 6 concludes by identifying current frontiers of the literature.

2. Foundations and Motivation

Empirical sorting models are motivated by a long-standing question. How can we estimate the demand for public goods that are not explicitly traded in formal markets? Early research sought to estimate demand by simply regressing the expenditures made for local public services on the characteristics of voters. Tiebout (1956) recognized that households “vote with their feet”. These migration patterns can bias reduced-form estimation. Mitigating the bias requires using knowledge of the sorting process to develop instrumental variables that isolate the underlying structure from data on the observed equilibrium outcomes. Formal models of the sorting process were developed using the characteristics approach to consumer theory (Kelvin J. Lancaster 1966; William M. Gorman 1980). The remainder of this section describes these foundations of the new equilibrium sorting literature.

2.1 A Reduced-Form Approach to Estimating the Demand for Public Goods

Theodore C. Bergstrom and Robert P. Goodman (1973) were among the first to propose a strategy for estimating the demand for local public goods. They envisioned an urban landscape in which the level of public goods supplied by each community is determined by that community’s median voter. Assuming the median voter also has the median level of income, the demand for a public good can be estimated by simply regressing actual expenditures on the local government services that are hypothesized to be associated with the specific public goods of interest. In equation (1) A represents one such set of expenditures, y_{med} represents the incomes of the median households, and τ_{med} represents the marginal tax rates faced by those households in each of several communities,

$$(1) \quad \ln A = \beta_0 + \beta_1 \ln y_{med} + \beta_2 \ln \tau_{med} + \sum_k \beta_k d_{k,med} + u_{med} .$$

The d_k 's describe the median household’s demographic characteristics. The simplicity of Bergstrom and Goodman’s estimator inspired numerous applications to community level data, as well as a microeconomic extension of the model to individual survey data (Bergstrom, Daniel L. Rubinfeld and Perry Shapiro 1982).

The problem with estimation of (1) is that it ignores the sorting process that underlies

equilibrium in the market for housing. If households choose where to live based, in part, on their preferences for public goods, the community selection mechanism implies the error, u_{med} , can be correlated with the independent variables affecting sorting. This correlation can bias estimates of the price and income elasticities. Gerald S. Goldstein and Mark V. Pauly (1981) labeled this problem “Tiebout bias” after Tiebout’s (1956) conceptual model of local public goods provision.¹

To illustrate this bias, consider an example from Rubinfeld, Shapiro and Judith Roberts (1987). Suppose household i maximizes its utility by locating in one of a discrete set of J communities based, in part, on its preferences for public goods,

$$(2) \quad j = \max_{j \in J} V(A_j, \tau_j, y_{ij}, d_i).$$

Then the estimating equation from Bergstrom and Goodman’s (1973) model can be rewritten for an individual observation as

$$(3) \quad \ln A_{i,j} = \beta_0 + \beta_1 \ln y_{i,j} + \beta_2 \ln \tau_{i,j} + \sum_k \beta_k d_{k,i,j} + u_{i,j}.$$

Reformulating the problem in terms of individual behavior allows us to interpret the econometric error term as a function of unobserved preferences. In this context, preference based sorting presents a simultaneity problem. A household’s income will depend on its location choice if communities differ in the job opportunities they provide. A community’s marginal tax rate will depend on the composition of its residents if tax rates are determined by voting. If income and taxes depend on location choices that are driven, in part, by unobserved preferences, $u_{i,j}$ will be correlated with $y_{i,j}$ and $\tau_{i,j}$, biasing OLS estimation of (3).

Rubinfeld, Shapiro, and Roberts (1987) propose a two-step selection model that has the potential to provide consistent estimates of the demand for a public good in the presence of Tiebout bias. While the logic behind their estimator is straightforward, there are two major hurdles to implementation. First, their strategy requires “good” instruments for the endogenous variables in (3). This requirement poses a challenge because the validity of any potential instru-

¹ Tiebout (1956) envisioned freely mobile households migrating across communities based on their preferences for the public goods provided by those communities. This type of sorting behavior poses a problem for OLS estimation of (1) regardless of whether the data describe the median household in each community or a random sample of households.

ment depends on the ways in which the sorting of heterogeneous households influences the properties of equilibria. Second, in most applications we must acknowledge that the model is incomplete. Omitted variables could easily differ across communities in ways that confound the identification of key parameters separately from Tiebout bias. Thus, the task of developing consistent estimates of the demand for a public good requires an empirical strategy for mitigating omitted variable bias together with knowledge of the sorting process.

Beginning with Sandra Black (1999), numerous studies have demonstrated that omitted variables and endogeneity of included variables can be of first-order importance for linear regressions using housing data, and have suggested innovative solutions (e.g. Lucas Davis 2004, Kenneth Y. Chay and Michael Greenstone 2005, Jaren C. Pope 2008). All models describing a behavioral process face these types of problems. So the choice of a structural or a reduced form approach does not in itself imply problems will arise or be solved. The modeling strategy creates different information needs as well as different opportunities to use information. Equilibrium sorting models provide an explicit representation for the sorting mechanisms that generate Tiebout bias, which can also be used to formalize the assumptions needed to ensure the validity of strategies for addressing omitted variables.

2.2 *A Model of Household Location Choice*

Equilibrium models of Tiebout sorting begin with a simple premise: the amount and character of housing and public goods varies across an urban landscape, and each household selects its preferred bundle of public and private goods given its income and the relative prices involved. Every household pays for its location choice through the price of housing. Working households may also pay indirectly through the wages they earn. As noted, a household's location choice reveals its preferences for a local public good.

Assume the urban landscape consists of $n = 1, \dots, N$ houses that can be divided into communities. Each house can be defined by a bundle of housing characteristics and amenities. h_n is a vector of structural characteristics that fully describe the private good component of an individual house. For example, h_n could include the number of bedrooms, the number of bathrooms, square feet, and lot size. g_j denotes a vector of amenities conveyed to every household

in community j . It may include local public goods such as school quality, urban and environmental services (such as crime rates, and air quality), and variables describing the demographic composition of the community (such as race and educational attainment). The term “amenities” is used to refer to any of these non-market goods and services.

A household’s utility depends on the characteristics of housing and amenities at its location and on its consumption of a composite numeraire private good, b . Households are heterogeneous. They differ in unobservable features of their preferences (α) and in observable factors such as their demographic characteristics (d). Let the population of households be indexed from $i = 1, \dots, I$. The utility obtained by household i from house n in community j can be represented as: $U(b, h_n, g_j; \alpha_i, d_i)$.²

Each household is assumed to choose a location and a quantity of b that maximize its utility subject to a budget constraint:

$$(4) \quad \max_{n \in j, b} U(b, h_n, g_j; \alpha_i, d_i) \text{ subject to } y_{i,j} = b + P_{n \in j}.$$

In the budget constraint, the price of the numeraire is normalized to one and $P_{n \in j}$ represents annualized expenditures on house n in community j ; in other words, $P_{n \in j}$ is the after-tax cost of occupying a single house for one year. $y_{i,j}$ is the household’s total annual income. The j subscript on income recognizes that, in general, income is endogenous to location choice.³

The specification in (4) implicitly removes three potential sources of “friction” from the problem. Households are assumed to have full information, to be freely mobile, and to make location choices within a discrimination-free environment. These three assumptions are maintained throughout the literature with very few exceptions.

Full information implies that households share the same objective assessment of the spatial landscape. In other words, every household fully observes the joint distribution of h , g , and P . It is also standard to assume that the government observes any variables being taxed (such as income or property values). In contrast, α_i is usually treated as private information. It is never

² A household may contain many members with different demographic characteristics and preferences, but it is treated as an indivisible economic agent.

³ Wages will be endogenous if heterogeneous workers sort across a landscape with spatial variation in the composition of labor demand. Even if employment locations are fixed, income may be endogenous to housing decisions because of commuting costs.

observed by the econometrician, and it is rarely assumed to be observed directly by other households. Exceptions to the later include models in which voters anticipate the consequences of their votes for housing prices and migration (e.g. Epple and Thomas Romer 1991) or homeowners anticipate the loss of open space (Walsh 2007).

Free mobility implies that households are free to move within the geographic region defined as the choice set. Theoretical and empirical models usually abstract from the physical and psychological costs of moving.⁴ However, it is sometimes recognized that moving to a new house may increase commuting costs or require a worker to simultaneously move a new job with a lower wage (e.g. Bayer, Robert McMillan, and Kim Rueben 2004; Kuminoff 2010).

Finally, there is assumed to be *no discrimination* in the sense that every household faces the same schedule of housing prices. This assumption does not rule out the possibility that prices could vary based on demographic features of households in a neighborhood. These could be proportions of households in specific ethnic or racial groups or fractions with school age children, for example. Sorting based on demographic features of neighbors is cited as one reason why some neighborhoods are effectively segregated by race (John Yinger 1976; David M. Cutler, Edward L. Glaeser, and Jacob L. Vigdor 1999).

Under the assumptions of free mobility, full information, and no discrimination, equilibrium is achieved when every household occupies its utility-maximizing location and nobody wants to move, given housing prices, housing characteristics, wages, tax rates, and amenity levels. Theoretical models investigate the existence and uniqueness of equilibria and their implications for equity and efficiency. Empirical models use the properties of equilibria to infer preferences for amenities from the observable characteristics of households and their location choices. This is done, in part, by adding a parametric specification for the neighborhood selection process to explicitly model the mechanisms that would generate Tiebout bias in a reduced form regression. Finally, because empirical models characterize equilibrium for an entire market, they can be used to assess market responses to policies that would make large changes to amenities.

⁴ In empirical work this is largely due to the absence of data on moving costs. Bayer, Nathaniel Keohane, and Timmins (2009) demonstrate that some moving costs can be estimated using related information, such as the location of an individual's hometown. Bayer et al. (2011) refine this strategy by using additional information about the timing of moves to estimate the psychological component of moving costs.

3. *Equilibrium Sorting Theory*

The theoretical literature that followed Tiebout's early work focused on formalizing his conceptual model, proving the existence of an equilibrium in which no household would be better off by moving, and extending his framework to include peer effects and other forms of social interaction within communities. Most of this research has focused on the demand side of the market, refining the characterization of household preferences and how equilibrium sorting influences *both* prices and endogenous amenities. Land developers are not explicitly represented. Instead, housing supply is usually treated as fixed or represented in a way that allows simple calibration, such as using a constant elasticity supply function.⁵ Modeling supply is a clear area for future research and we return to this topic in section 6. Part 1 of this section outlines how sorting equilibrium was described with increased generality. Part 2 adds social interactions and equity considerations. Part 3 describes how the hedonic price function relates to this framework.

3.1 *Equilibrium Stratification Patterns*

For heuristic purposes, a household's choice process can be depicted as a two-stage problem. Each household first determines the optimal quantities of housing and numeraire in each of a finite number of communities and then chooses the particular community that maximizes its utility.⁶ The first stage is

$$(5) \quad \max_{\bar{h}, b} U(b, \bar{h}, g_j, \alpha) \text{ subject to } y = b + p_j \cdot \bar{h}.$$

Theoretical models usually treat housing as a homogeneous commodity that can be consumed in continuous quantities at a constant price per unit. This is represented in (5) by defining p_j as the annualized per-unit price of housing in community j and \bar{h} as the quantity of housing consumed. The bar on \bar{h} does not imply the quantity of housing consumed is fixed; the quantity consumed may vary across households according to their income, preferences, and location. The bar is meant to highlight the process of constructing an index number to describe a homogeneous unit of housing. When h was used earlier (in equation(4)) it included the vector of specific structural

⁵ There are notable exceptions such as Walsh (2007) who includes a land supply model in characterizing the effects of policies to protect open space on the equilibrium supply of these amenities.

⁶ We have dropped the d term describing the individual's observable demographic characteristics at this stage to simplify notation.

characteristics (e.g. bedrooms, bathrooms, square feet) that would in principle contribute to the construction of the \bar{h} index. This distinction is important because it is one of the features that contributes to the differences between the pure characteristics and the random utility models, as we explain further below.

Assuming households can purchase any quantity of housing at the market price in each community, the amount of housing is “optimized out” of the problem and preferences can be restated using the indirect utility function in (6).

$$(6) \quad V(g, p, \alpha, y) = U[g, \bar{h}(g, p, \alpha, y), y - p \cdot \bar{h}(g, p, \alpha, y), \alpha].$$

Each household will choose the community that maximizes its well-being, given income and prices. A sorting equilibrium is achieved when every household has chosen its utility-maximizing community and nobody wants to move, given housing prices and public goods.

Bryan Ellickson (1971) first characterized the restrictions on preferences that induce sorting behavior. Three features of his model form the basis for most of the early studies evaluating the properties of the sorting equilibrium. First, he assumed that provision of public goods in community j could be represented by a 1-dimensional measure, \bar{g}_j , which represents an index of a set of public goods in that community. Second, he assumed that households have homogeneous preferences ($\alpha_i = \dots = \alpha_l$) and therefore differ only in their income. Given the first two assumptions, he imposed the restriction that indifference curves in the (\bar{g}, p) plane are strictly increasing in income. This, he reasoned, would support a sorting equilibrium in which households are perfectly stratified across communities by income. Figure 1 uses a two-community example to illustrate the idea. Household i is exactly indifferent between the two communities. Any household with lower income, such as household $i - 1$, will always prefer the cheaper community because indifference curves cannot cross more than once. Conversely, any household with higher income, such as $i + 1$, will always prefer the more expensive community.

Using the three restrictions from Ellickson’s paper, Frank Westoff (1977) proved that a sorting equilibrium exists in a model where households in each community vote to determine public goods provision and community-specific tax rates. Epple, Radu Filimon, and Romer (1984, 1993) extended Westoff’s model to include a market for housing that must clear within each community. Finally, Epple and Romer (1991) generalized the model further to allow voters

to anticipate the consequences of their votes for housing prices and migration. While these models help to formalize Tiebout's theory, they do a poor job of reproducing the outcomes observed from actual sorting behavior. To see why, let the J communities be ordered by their quality of public goods provision: $\bar{g}_1 < \bar{g}_2 < \dots < \bar{g}_J$. The key restriction in the early theoretical literature—Ellickson's single-crossing restriction—implies that households are partitioned across communities by income, as illustrated in figure 2. In the figure, every household in community 1 has lower income than every household in community 2, and so on. By contrast, actual community-specific income distributions overlap substantially.

One explanation for why households do not perfectly stratify by income is that they differ in their tastes for public goods. Recognizing this point, Epple and Glen J. Platt (1998) extended the Epple-Romer model to allow households to differ in a heterogeneous parameter that represents their preferences for composite provision of public goods relative to private goods. In this case, equilibrium is characterized by a more general single-crossing restriction. To formalize the restriction, equation (7) shows the slope of an “indirect indifference curve” in (\bar{g}, p) space.

$$(7) \quad M(\bar{g}, p, \alpha, y) = \left\langle \frac{dp}{d\bar{g}} \middle| V = \bar{V} \right\rangle = - \frac{\partial V(\bar{g}, p, \alpha, y) / \partial \bar{g}}{\partial V(\bar{g}, p, \alpha, y) / \partial p}.$$

Assuming M is monotonically increasing in $(y | \alpha)$ and $(\alpha | y)$, indifference curves in the (\bar{g}, p) plane satisfy single crossing. This 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, $\bar{h}(\bar{g}, p, \alpha, y)$.

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

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.⁷

⁷ 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 H. Spencer Banzhaf (2004, 2007) and Raymond B. Palmquist (2005a) for details. David S. Bullock and Nicholas Minot (2006) have demonstrated that the Willig condition is sufficient but not necessary for identifying willingness to pay for changes in non-market goods with weak complementarity.

The single crossing condition implies that, in equilibrium, three properties characterize sorting by each household type: *boundary indifference*, *stratification*, and *increasing bundles* (Epple and Platt 1998). Without loss of generality, let the J communities be ordered according to the index of public goods, $\bar{g}_1 < \dots < \bar{g}_J$. *Boundary indifference* requires a household on the “border” between two communities in (α, y) space to be indifferent between those communities. Equation (9) defines the set of border individuals. It must hold for all $j = 1, \dots, J - 1$.

$$(9) \quad \left\{ (\alpha, y): V(\bar{g}_j, p_j, \alpha, y) = V(\bar{g}_{j+1}, p_{j+1}, \alpha, y) \right\}.$$

The *increasing bundles* property requires that for any two communities in the ordering, $(j, j + 1)$ equation (10) must hold.

$$(10) \quad y_{j+1}(\alpha) > y_j(\alpha) \Rightarrow p_{j+1} > p_j \quad \text{and} \quad \bar{g}_{j+1} > \bar{g}_j.$$

That is, the ranking of communities by public goods provision must match the ranking by price. The third property, *stratification*, requires that households of each type are stratified across the J ordered locations by $(\alpha | y)$ and by $(y | \alpha)$, as defined in (11).

$$(11) \quad \begin{aligned} & (y_{j-1} | \alpha) < (y_j | \alpha) < (y_{j+1} | \alpha) \\ & \text{and} \\ & (\alpha_{j-1} | y) < (\alpha_j | y) < (\alpha_{j+1} | y) \end{aligned}.$$

Figure 3 illustrates the implied partition of households into communities in (y, α) space. Conditional on preferences, wealthier households always choose to live in communities with more public goods. Likewise, conditional on income, households with stronger preferences choose communities with more public goods. This two-dimensional stratification is consistent with Tiebout’s (1956) reasoning and can be reconciled with differences in observed income distributions across communities.

While *boundary indifference*, *stratification*, and *increasing bundles* are necessary for a sorting equilibrium to exist, they are not sufficient. Any sorting equilibrium must also be characterized by a vector of housing prices and a vector of public goods such that no household could increase its utility by moving. The development of a general existence proof is complicated by preference heterogeneity. Epple and Platt (1998) rely on a Cobb-Douglas specification for

utility and use a numerical example to demonstrate that a sorting equilibrium exists. Sieg et al. (2004) reinforce their finding by demonstrating existence numerically using a constant elasticity-of-substitution (CES) specification for utility and Kuminoff (2009) allows households to differ in their relative preferences for multiple public goods.

Thus far, we have characterized the properties of equilibria that arise from the two-way interaction between households and the community-level provision of local public goods. Existing residents of a community are assumed to be indifferent to households that migrate into the area unless their actions change housing prices or alter voting outcomes. This assumption overlooks the evidence that people care about the demographic characteristics of their neighbors (e.g. Cutler, Glaeser, and Vigdor 1999). The demographic composition of the neighborhood may also affect the quality of endogenous public goods. For example, parents appear to judge the quality of public schools by how students perform on standardized tests (Jesse M. Rothstein 2006) and the best predictors of a student's performance tend to be the demographic characteristics of her parents (particularly income and education) and the performances of her peers (Eric A. Hanushek 2003). Since endogenous public goods are a key challenge for empirical estimation, it is important to understand how they may arise.

3.2 Social Interactions, Equity, and Efficiency

Charles A.M. de Bartolome (1990) was the first to build social interactions into a model of residential choice. He depicts two types of households (low skill and high skill) sorting across two communities according to their differentiated preferences for a single public good—school quality—which is increasing in expenditures and household skill.⁸ This simple framework leads to two insights that are robust to many extensions of the model: (i) social interactions can produce multiple equilibria, and (ii) some of the equilibria are inefficient. The nature of the equilibrium depends on the strength of the peer group effect. If peer effects have little impact on school quality, then a single-crossing condition is sufficient to guarantee that household types will be segregated by skill. As peer effects grow in importance, low skill households have a stronger incentive to move to the high skill community. This can lead to an “integrated” equilibrium in

⁸ While De Bartolome (1990) was the first to characterize the consequences of social interactions for sorting equilibria, Tiebout (1956) recognized their potential role in location choice. In footnote 12, he states: “*Not only is the consumer-voter concerned with economic patterns, but he desires, for example, to associate with ‘nice’ people.*”

which both communities contain both types, and each community provides a different level of education. Interestingly, the social interactions that underlie this equilibrium also cause it to be inefficient. Social interactions generate externalities. In De Bartolome's (1990) model, migrants do not internalize the effects of their location choices on the residents of the destination communities. This effect of their relocation is the underlying source of inefficiency⁹.

Thomas J. Nechyba (1999) extends De Bartolome's model to consider I household types sorting among J school districts under the added assumption that each district contains multiple housing types of varying quality. In a calibrated version of the integrated equilibrium, wealthier districts have higher quality schools but there is some overlap in district income distributions. The policy intervention is a voucher program. Nechyba introduces private schools as a club good and evaluates the impact of partial tuition vouchers on equilibrium outcomes. Parents in wealthy districts are the first to react. They move to poor districts, where housing is cheap, using vouchers to send their children to private schools. Because peer quality is increasing in parental income, public school quality declines in the wealthy districts. Average peer quality also declines in poor districts because the wealthy immigrants displace the prior occupants of the most expensive houses (who use the capital gains from selling their houses to move to wealthy districts). As a result, the poorest public schools lose their best students. However, this decrease in peer quality is at least partly offset by increased education expenditures if districts vote on spending. Per capita expenditures increase in poor districts due to smaller class sizes and increased property tax revenues. Thus, Nechyba's analysis suggests voucher programs may reduce inequality between rich and poor districts due, in part, to migration, voting, and peer effects.

Roland Bénabou (1993) demonstrates that by influencing the production of education, social interactions and stratification in the housing market can also have general equilibrium implications for efficiency and growth. This result follows when the economy includes a production sector with complementarity between high and low skill labor.¹⁰ Not only does the incentive to segregate raise the cost of education in low skill communities; it also increases unemployment, which decreases production, exacerbating the inefficiency from stratification.

¹⁰ He also moves to a representative agent framework where homogeneous households choose whether to be unemployed or to invest in education to obtain varying degrees of skill.

These effects can be persistent. In a dynamic version of the model, Steven N. Durlauf (1996) demonstrates that short run stratification can have long-term consequences for inequality and growth. Parents who had the misfortune to be born into a poor community may be unable to raise school quality enough for their children to obtain higher paying jobs. One potential solution to this poverty trap is to equalize expenditures on education across districts. Bénabou (1996a) models the implications of this approach. The net benefits hinge on an intergenerational tradeoff between the short run cost of constraining expenditures on education and the long run benefit of reducing inefficiency from stratification. Most of the key results from these “general equilibrium” sorting models are provided in Bénabou (1996b) who observes that minor differences in preferences can create a “tipping” effect that leads to severe stratification by income.

The tipping effect also helps to explain the persistence of racial segregation. In Rajiv Sethi and Rohini Somathathan’s (2004) model of sorting with social interactions, households prefer integrated communities to segregated ones but, if forced to choose between two racially segregated communities, would prefer to live in the one occupied by their own race. A single crossing condition ensures that, all else constant, households will be stratified by income. To illustrate how the two effects can combine to sustain segregation, consider an initial equilibrium where households are effectively stratified by race due to a large black-white income gap, but would prefer to be integrated. As the income gap narrows, a rich black household living in the predominantly black community has *less* of an incentive to move to the predominantly white community because the white community has become less affluent in relative terms.

The single-crossing assumption is usually maintained to help characterize the properties of equilibria. However, this assumption is not necessary to guarantee that equilibria exist. Nechyba (1997) proves existence after imposing exogenous community boundaries and endowing households with fixed quantities of housing. Bayer and Timmins (2005) also prove existence in a model with social interactions by smoothing the preference function. They add an idiosyncratic shock to utility and, in lieu of endogenous prices, they allow utility to depend on the share of households in each community.¹¹ This endogenously determined share can be interpreted as a (negative) congestion effect or a (positive) agglomeration effect. Importantly, the equilibrium is

¹¹ With an estimate of the supply curve for housing, it is a simple matter to go from this specification to one that includes the price of housing directly in utility.

shown to be unique in the case of congestion. In the case of agglomeration, whether the equilibrium is unique depends on the strength of preferences for the endogenous amenity.¹²

3.3 The Equilibrium Hedonic Price Function

Hedonic price functions provide an alternative way to characterize sorting equilibria. If utility is continuously differentiable, monotonic in the numeraire, and Lipschitz continuous, Bajari and Benkard (2005) prove that, in equilibrium, the price of a differentiated product will be a function of its characteristics. Thus, under mild restrictions on preferences, the equilibrium price of an individual home can be expressed as a function of its structural characteristics and the amenities it provides: $P_{n \in j} = P(h_n, g_j)$. Empirical hedonic models were first popularized through Zvi Griliches's (1961) research using hedonic price functions as a basis for quality adjustments to price indices for automobiles.¹³ Rosen (1974) later strengthened the economic foundations of the method, arguing that the hedonic price function can be interpreted as describing an equilibrium relationship as a continuous array of differentiated varieties of a product at a single point in time.¹⁴ Rosen recognized that applying the model to housing offered a special case of Tiebout's (1956) original model. In the special case with no economies of scale in producing public goods, Tiebout (1956) suggests that households would choose communities that *exactly* match their preferences, effectively making each household its own local government.¹⁵ Likewise, Rosen (1974) observes that the hedonic price function reflects equilibrium stratification patterns that mirror those in Tiebout's work.¹⁶

¹² These results can be interpreted using our earlier description of the single crossing condition. It restricts how the virtual price of public goods per unit of housing change with income. Holding the quantity of housing constant in Nechyba's case allows for a wider set of patterns of virtual price change with income. Bayer and Timmins dampen the effects of income on the value of public goods by introducing the offsetting congestion effect.

¹³ While empirical examples of hedonic modeling date back to Frederick V. Waugh's (1929) PhD Thesis, the first contributions to an underlying theory were made by A.D. Roy (1951) and Jan Tinbergen (1959). Their work focused on the market for labor. Roy (1951) argued that the equilibrium distribution of wages would reflect the underlying distributions of preferences and skills held by all of the producers and consumers in a market and Tinbergen (1959) provided the first analytical demonstration of this logic. However, their contributions to hedonic theory were not widely recognized until much later.

¹⁴ There is an important dichotomy in the literature that distinguishes Griliches (1961) and Jack E. Triplett's (1969) interests from those that followed Rosen (1974). Triplett's (1983) account of the development of the hedonic price function for price indexes as models without theory offers an interesting contrast to the post Rosen (1974) literature. His focus is the role of the price function in developing price and quantity indexes for quality differentiated goods rather than on the interpretation of its derivatives.

¹⁵ From page 421 of Tiebout (1956), "...the consumer-voters will move to that community which *exactly* satisfies their preferences. ...this may reduce the solution of the problem to the trite one of making each person his own municipal government."

¹⁶ From page 40 of Rosen (1974), "...a clear consequence of the model is that there are natural tendencies toward market segmentation, in the sense that consumers with similar value functions purchase products with similar specifications. ...In fact,

The hedonic property value model illustrates how a sorting equilibrium can provide information about the demand for amenities. Consider a single amenity, g_1 . Partially differentiating the equilibrium price function with respect to g_1 provides an estimate of the marginal price function for g_1 :

$$(12) \quad P_1 = \frac{\partial P(h, g)}{\partial g_1}.$$

P_1 is the marginal contribution of g_1 to the price of housing given the current level of g_1 and levels of the other characteristics.

Because households are assumed to face a continuum of choices, the first order conditions for the utility maximization problem for g_1 can be expressed as:

$$(13) \quad \frac{\partial P(h, g)}{\partial g_1} = \frac{\partial U / \partial g_1}{\partial U / \partial b} \equiv D(g_1; g_{-1}, h, \alpha, y).$$

The first equality implies that households will choose a location that provides them a level for g_1 where their marginal willingness-to-pay for an additional unit exactly equals its marginal implicit price. Assuming the marginal utility of income is constant, the identity observes that, as g_1 varies, the marginal rate of substitution defines the household's inverse demand curve, conditional on all other amenities (g_{-1}) and housing characteristics.

Figure 4 shows bid functions for housing in the g_1 dimension for two different households.¹⁷ Each household will select the quantity of g_1 where its bid function is tangent to the hedonic price function. In the figure, the two households purchase houses that are identical except in their provision of the public good. Household 1 spends $\$_1$ on a house that provides $g_{1,1}$ units of the public good and household 2 spends $\$_2$ on a house with $g_{1,2}$.

If markets are in equilibrium, evaluating P_1 at a household's chosen level of g_1 will return that household's marginal willingness-to-pay (MWTP) for g . Combining this information with the level of g_1 at a household's location identifies exactly one point on that household's

the above specification is very similar in spirit to Tiebout's (1956) analysis of the implicit market for neighborhoods, local public goods being the 'characteristics' in this case."

¹⁷ The bid functions express each household's willingness-to-pay for housing as a function of the amenity, given the household's preferences, income, levels of all the other characteristics, and the utility attained.

inverse demand curve. In figure 5, household 1's inverse demand (D_1) intersects P_1 at the point where its MWTP exactly equals the marginal price for an extra unit of g_1 . While MWTP is identified by the gradient of the price function, inverse demand curves are generally not. An infinite number of demand curves could pass through the points defined by $(MWTP_1, g_{1,1})$ and $(MWTP_2, g_{1,2})$.

Moreover, if households are not free to choose continuous quantities of each amenity, the price function need not identify MWTP. There is an important distinction between the price function as a description of an equilibrium and the information it conveys about preferences. The hedonic price function can be used to describe a sorting equilibrium, regardless of whether its characteristics are discrete or continuous (Bajari and Benkard 2005). However, if at least one characteristic is discrete, the first order condition in (13) will not adequately characterize equilibrium behavior. Moreover, there is not a specific link between the marginal price and the MWTP. This qualification can be important in practice because many amenities do vary discretely. When amenities vary discretely across the spatial landscape, utility maximization is characterized by the set of inequalities in (14) rather than the first order conditions in (13).

$$(14) \quad U(y_{i,j} - P_{n \in j}, h_n, g_j; \alpha_i) \geq U(y_{i,k} - P_{m \in k}, h_m, g_k; \alpha_i), \quad \forall m, k.$$

The equation simply says that if household i chooses house n in community j , it is because that location provides it with at least as much utility as any other alternative in its budget set.¹⁸

Econometric approaches to demand estimation can be divided into two frameworks: hedonic models that exploit the first-order conditions in (13) and sorting models that exploit the inequalities in (14) together with the equilibrium price function (see Palmquist (2005b) for a review of empirical hedonic models). Both frameworks assume that households “pay” for amenities through the price of housing and then use data on housing prices and spatial variation in amenities to infer the demand. This strategy presents a fundamental identification problem. Since each household is typically observed making a single housing purchase it is possible to

¹⁸ Comparing a bundle of goods that was purchased with bundles that could have been purchased, but were not, can serve to identify bounds on a consumer's indifference curves (Paul A. Samuelson 1948). More precisely, assuming preferences are monotonic and convex, it is possible to recover bounds on the set of indifference curves that would be consistent with utility maximizing behavior (Hal R. Varian 1982). To identify a household's indifference curve within these bounds, however, the analyst must be willing to impose more structure on preferences (Kuminoff 2009).

identify, at most, one point on that household's demand curve. In order to recover the entire demand curve, or to use results to evaluate any non-marginal change, the analyst must add information about the structure of preferences in an equivalent way to what has been done with the models used for equilibrium sorting estimation.

4. *Equilibrium Sorting Estimation*

Most applications of sorting models are linked to one of the three predominant frameworks developed by Epple and Sieg (1999), by Bayer, McMillan, and Rueben (2004), and by Nechyba (1999, 2000) and Ferreyra (2007). Each framework differs in how it defines the choice process, how it depicts household preferences, how it uses the properties of equilibria to develop instrumental variables, and how it approaches econometric estimation. After assessing the strengths and weakness of each estimation framework, we discuss calibration as an alternative strategy. The section concludes with a brief summary of the tradeoffs that have arisen with efforts to describe sorting equilibria using quasi-experimental research designs.

4.1 *The Pure Characteristics Sorting Model (hereafter labeled as PC)*

Epple and Sieg (1999) provided the first illustration of how the properties of a sorting equilibrium can be used to recover households' preferences for amenities. Their structural model parallels the theoretical literature on equilibrium stratification discussed in section 3.1 (Ellickson 1971; Westoff 1977; Epple, Filimon, and Romer 1993; and especially Epple and Platt 1998). Housing is treated as a homogeneous commodity that can be consumed at a constant (community-specific) price. With this assumption, the price of housing reflects the cost of consuming the amenities provided by each community.

In practice, housing is not homogenous. Its structural characteristics 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 preferences as part of a separable sub-function that is homogeneous of degree one. This restriction implies the quantity index for the amount of housing can be defined regardless of the level of amenities and of the numeraire. The equilibrium locus of housing expenditures can then be

expressed as the product of a price index and a quantity index: $P_{n \in j} = p(g_j) \cdot \bar{h}(h_n)$. This is the basis for connecting to structural attributes and the former the mechanism describing how community specific public goods affect prices for a homogenous unit of housing.¹⁹ Taking logs produces the hedonic property value model in (15).

$$(15) \quad \ln(P_{n \in j}) = \ln[\bar{h}(h_n)] + \ln[p(g_j)].$$

Given a parametric form for the quantity index and micro data on housing transaction prices and their structural characteristics, the price of housing in each community can be estimated as a community-specific fixed effect: p_1, \dots, p_j .

The index of structural characteristics is optimized out of the indirect utility function under the standard hedonic assumption that households are free to choose continuous levels of structural characteristics, no matter where they choose to live. Thus, the PC logic depicts a discrete-continuous optimization problem where households choose one of a discrete number of communities and then, conditional on that selection, a continuous quantity of housing.

Equation (16) shows the CES indirect utility function for household i in community j . The first term represents utility from community-specific amenities. \bar{g} is a linear index of K amenities, all but one of which are observable. The K^{th} amenity ($g_{K,j} = \bar{\xi}_j$) represents an index of all the community-specific attributes that are not observed by the econometrician.

$$(16) \quad V_{i,j} = \left\{ \alpha_i (\bar{g}_j)^\rho + \left[\exp\left(\frac{y_i^{1-\nu} - 1}{1-\nu}\right) \exp\left(-\frac{\beta(p_j^{\eta+1}) - 1}{1+\eta}\right) \right]^\rho \right\}^{\frac{1}{\rho}},$$

$$\text{where } \bar{g}_j = \gamma_1 g_{1,j} + \gamma_2 g_{2,j} \dots + \gamma_{K-1} g_{K-1,j} + \gamma_K \bar{\xi}_j, \quad F(\alpha, y) \sim \text{lognormal}.$$

Notice that the weights in the public good index are constant ($\gamma_1, \dots, \gamma_K$). This restriction implies households agree on the ranking of communities by \bar{g} . Households can differ in the strength of their preferences for amenities relative to private goods through the α_i term. To implement the model empirically, the joint distribution of α with income is assumed to be

¹⁹ Another interpretation of this result is that it offers a sufficient set of conditions for defining the price indexes for goods with difficult to measure quality attributes. As a result, it helps to connect the equilibrium descriptions of hedonic price functions with the price index descriptions often associated with Triplett (1983).

lognormal.

The second term in the CES function represents utility from private goods. Applying Roy's Identity yields the demand function for the private good component of housing:²⁰

$$(17) \quad \bar{h}_{i,j} = \beta p_j^\eta y_i^\nu.$$

The price and income elasticities of the demand for housing are represented by η and ν , β is a housing demand intercept, and ρ reflects substitution between public and private goods.²¹ Given the expected signs for the housing demand parameters, the single crossing condition implies $\rho < 0$. This condition allows for a test of whether the parameter estimates are consistent with the assumptions used to characterize the equilibrium conditions and to use the model for computing changes in equilibrium prices in response to exogenous policy changes.²²

Most applications of the PC model have followed Epple and Sieg and used a variation of their indirect utility function, which does not allow for idiosyncratic shocks associated with each choice $(\varepsilon_{i,1}, \dots, \varepsilon_{i,J})$. As a result the model must maintain that all households agree the communities differ only in the K characteristics that comprise the \bar{g}_j index.²³

In their application of the PC approach to the Boston Metropolitan Area, Epple and Sieg (1999) define each community as a school district. Using data on housing prices, household income, school quality, and crime rates in 92 school districts, they estimate the parameters of the model in two stages. First, they use the *stratification* property from (11) to express the quantiles of the income distributions in each community as a function of the structural parameters. A subset of these parameters can be identified by minimizing the distance between predicted and observed income quantiles. In the second stage, they use the *increasing bundles* and *boundary indifference* properties (9)-(10) to develop a nonlinear instrumental variables estimator to identify and estimate the remaining parameters, including a composite unobserved public good in each community $(\bar{\xi}_1, \dots, \bar{\xi}_J)$.

²⁰ The constant elasticity of demand function describing housing is consistent with the sub-function in (16). There is no direct utility function that corresponds to this constant demand elasticity form.

²¹ Specifically, the elasticity of substitution is defined as $\sigma = 1/(1 - \rho)$.

²² A limitation of this test is that the estimation algorithm maintains the stratification and increasing bundles properties. Thus, the power of the test depends on the domain of values for ρ capable of violating the single crossing property while simultaneously satisfying stratification and increasing bundles.

²³ The PC label stems from Berry and Pakes (2007), who labeled this property the "pure characteristics" approach to modeling the demand for a differentiated good to distinguish it from "random utility models" that assume idiosyncratic taste shocks.

Instruments are needed in this second-stage estimator because equilibrium prices (p) and amenities (\bar{g}) may be correlated with unobserved amenities captured by the community-specific fixed effects ($\bar{\xi}$). In this context $\bar{\xi}$ is the structural equivalent to an error term. It is important to consider the potential for any candidate instrument to distinguish the fixed effects from the coefficients on observed public goods. The *increasing bundles* property in (10) implies that p_j , \bar{g}_j , and $y_j(\alpha)$ will all follow the same ranking across communities. Epple and Sieg exploit this condition to define instruments based on functions of the income rank of each community. With the single crossing condition the identification logic does not preclude endogenous attributes. Rather the effects of income on the virtual price of the amenity per unit of housing must overcome any other effects of endogeneity. For the instruments to be valid, however, we must be willing to assume that unobserved amenities are of second order importance in the sorting process. While $\bar{\xi}$ may affect the *level* of income in each community, it must not affect the *ranking* of communities by income.

Sieg et al. (2004) refine this estimator by adding moment conditions based on the distribution of housing prices. They use a GMM approach to estimate all the parameters of the model simultaneously. Their application considers preferences for air quality, school quality, and crime rates in 93 school districts in the Los Angeles Air Basin. Other applications have focused on open space and access to recreation opportunities (JunJie Wu and Seong-Hoon Cho 2003, Walsh 2007). Stephen Calabrese et al. (2006) extend the PC logic and estimation framework to incorporate peer effects and majority rule voting on a property tax used to finance local public goods. All of these applications have found that $\rho < 0$, suggesting that observed sorting behavior is consistent with the single crossing condition.²⁴

One of the key limitations of the PC framework is the maintained assumption that all households agree on a common ranking of communities based on their provision of local public goods including exogenous environmental amenities. Relaxing this assumption is important because it is reasonable to expect households to evaluate the components of a vector of amenities quite differently. Epple, Michael Peress and Sieg (2010) suggest one could generalize the model

²⁴ As noted earlier, the stratification and increasing bundles properties are treated as maintained assumptions in the estimation frameworks used to test the single crossing condition.

by specifying different types of households based on observable attributes, and use a separate function to describe how location specific amenities and public goods affect each type's community choices.

4.2 *The Random Utility Sorting Model (hereafter labeled as RU)*

Bayer, McMillan, and Reuben (2004) develop a probabilistic approach to estimating household preferences. Following McFadden (1978), they begin by defining the object of choice as an individual house.²⁵ Then, similar to the PC approach, they assume each house is located within a community that provides a bundle of amenities. The resulting RU model is quite flexible in its treatment of preference heterogeneity. Households are allowed to differ in their relative preferences for multiple housing characteristics and amenities. In addition, each household is modeled as having an idiosyncratic “taste” for every choice alternative.

Equation (18) depicts household i 's utility from occupying house n in community j . For notational simplicity, i superscripts are used to index households and subscripts are used to index a partition of the structural preference parameters into four vectors that correspond to tastes for housing characteristics, amenities, commuting distance to the household's job location, and the price of housing, $\alpha_i = [\alpha_h^i, \alpha_g^i, \alpha_c^i, \alpha_p^i]$.

$$(18) \quad V_{i,n \in j} = \alpha_h^i h_n + \alpha_g^i g_j + \alpha_c^i c_{i,j} + \alpha_p^i p_n + \bar{\xi}_n + \varepsilon_n^i,$$

$$\text{where } \alpha_a^i = \alpha_{0,a} + \sum_{r=1}^R \alpha_{r,a} d_r^i, \quad \text{and } \varepsilon_j^i \sim \text{iid type I extreme value.}$$

A household's preferences are hypothesized to depend on its demographics. This characterization introduces heterogeneity linked to observables into the framework. It is modeled by decomposing preferences for each attribute into a common component, α_0 , and a set of terms that interact with the household's demographic characteristics, d^i . Finally, the unobservable (to the analyst) dimensions of heterogeneity, are assumed to be identically and independently distributed following a type I extreme value distribution.

Because idiosyncratic tastes follow the type I extreme value distribution, the probability of choosing a house (or a type of house) can be expressed as a function of the structural param-

²⁵ Subsequent versions of this paper and related research have aggregated the choice alternative into classes of houses for reasons described below.

ters (McFadden 1974). This relationship forms the basis for estimation. Bayer, McMillan, and Reuben adapt the instrumental variables estimator from Berry, Levinsohn and Pakes (1995) to recover values for the heterogeneous preference parameters that minimize the difference between predicted and observed location choices. The house-specific unobserved characteristics (ξ) are treated as structural error terms to be recovered during the estimation.

In their application to the San Francisco Metropolitan Area, Bayer, McMillan, and Reuben (2004) define each community as a census block group. They combine Census data on the prices and characteristics of houses within each community with data on a diverse set of community-specific amenities including crime rates, school quality, elevation, % white, % black, and average income. The endogeneity of housing prices and amenities is handled in estimation by developing two separate sets of instruments.

To address the correlation between observed and unobserved amenities in a similar analysis, Bayer, Ferreira and McMillan (2007) adapt the quasi-experimental research design from Sandra E. Black's (1999) reduced form analysis of the relationship between test scores and property values. First, they limit their sample to houses located less than a quarter mile from a school attendance zone boundary. Then they add fixed effects for each boundary to absorb the average effect of the omitted variables near each boundary. They argue that this allows them to isolate the variation in school quality that occurs across boundaries, as well as variation in the demographic composition of neighborhoods.²⁶ This strategy is an example of how a structural model can integrate insights from the literature exploiting quasi experimental research designs to define new instruments.

The instrument for price in community j builds on Berry, Levinsohn, and Pakes (1995) in that it is based on close substitutes and is calculated as a function of housing characteristics and exogenous amenities in all other communities. Bayer and Timmins (2007) demonstrate that, in equilibrium, the level of an endogenous attribute in community j will depend on the exogenous attributes of every other location. This relationship ensures that combinations of those exogenous attributes will be relevant instruments. The validity of these instruments follows from the maintained assumption that the utility from locating in community j is not directly affected by

²⁶ The notion that race and other demographic characteristics may also vary across school district boundaries is broadly consistent with Sethi and Somathathan's (2004) model of racial segregation.

the exogenous attributes of any other community.

A key requirement for establishing consistency and asymptotic normality of the estimator is that the number of households must grow large relative to the number of distinct housing “types”, where each type is defined by a unique bundle of structural characteristics and amenities. Bayer, McMillan, and Reuben address this requirement by assuming that all of the houses in a single Census block group with identical structural characteristics comprise a single type. Let N and I represent the total numbers of housing types and households. Consistency and asymptotic normality are defined as $N \rightarrow \infty$. The limit theorems developed by Berry, Oliver B. Linton, and Pakes (2004) can be invoked to establish consistency as long as $(N \log N)/I \rightarrow 0$. The additional restriction that N^2/I is bounded is sufficient to guarantee asymptotic normality at rate \sqrt{N} . Recognizing the importance of these results, subsequent applications explicitly aggregated choice alternative into classes of houses (e.g. Bayer, Ferreira and McMillan 2007; Klaiber and Phaneuf 2010b).

The basic RU model and identification logic from Bayer, McMillan, and Reuben (2004) have been adapted to various metro areas in the U.S. to estimate preferences for school quality, air quality, the demographic composition of a community, and open space (Bayer, Ferreira, and McMillan 2007; Tra 2010; Klaiber and Phaneuf 2010b) as well as in Switzerland where Kurt Schmidheiny (2006) evaluates the contribution of local income taxes to income stratification; in Brazil, with Timmins (2007) valuing climate amenities, and in Mumbai, India with Akie Takeuchi, Maureen Cropper, and Antonio Bento (2008) assessing slum improvement programs.

4.3 *The Calibrated Sorting Model (hereafter labeled as CS)*

Ferreira (2007) developed a third structural estimator based on the theoretical models and calibrated simulations developed by Nechyba (1997, 1999, 2000) and emphasizes school quality. Her approach relies on specific analytical assumptions rather than instruments to identify the endogenous variables. In effect she explicitly models the mechanisms that underlie the endogeneity. In addition to specifying household preferences, she specifies a production function for education. Households vote to tax themselves to fund public schools. She uses these relationships to estimate parameters describing household utility and the production of

education.²⁷ Each draw on the structural parameters used to predict housing choices, school choices, voting outcomes, tax rates, and expenditures on public education.

Compared to the PC and RU models Ferreyra's preference specification is simple. Households have identical Cobb-Douglas preferences over school quality (s), an index of all other community attributes (\bar{g}), and the composite numeraire (b),

$$(19) \quad U_{i,j} = \alpha_s \ln s_{i,j,t} + \alpha_{\bar{g}} \ln \bar{g}_j + (1 - \alpha_s - \alpha_{\bar{g}}) \ln b_{i,j} + \varepsilon_{i,j}.$$

The only explicit source of preference heterogeneity is an idiosyncratic shock for each individual and location, $\varepsilon_{i,j}$, which is assumed to follow the usual *iid* Type I extreme value distribution.

Households may differ in their *perceptions* about the quality of public schools relative to private Catholic schools. Equation (20) depicts the quality of two schooling types t (public, $t=1$ and private, $t=2$) in community j as perceived by household i ,

$$(20) \quad \begin{aligned} s_{i,j,1} &= q^\phi A^{1-\phi} && \text{for a public school, and} \\ s_{i,j,2} &= \alpha_C^i q^\phi A^{1-\phi} && \text{for a private Catholic school.} \end{aligned}$$

In this Cobb-Douglas production function ϕ is a technology parameter to be estimated, A is expenditures per student, and peer effects are denoted by q , which is set equal to the average income of households with children attending the school. Perceptions of school quality, α_C^i , are assumed to be uniformly distributed, with one distribution for Catholic households and another distribution for non-Catholics.

Amenities other than school quality are less important in Ferreyra's (2007) model. The composite index of non-school amenities, \bar{g}_j , is recovered through a preliminary hedonic regression of housing expenditures on their structural characteristics, community amenities, and a set of fixed effects for school districts. This strategy is essentially the same as the procedure suggested by Sieg et al. (2002) and illustrated in (15). The average predicted price for the bundle of remaining explanatory variables is used as a measure of \bar{g}_j .

Households in her model make three choices: (i) where to live, (ii) whether to send their

²⁷ On a given iteration of her minimum distance estimator, Ferreyra uses the structural parameters to simulate equilibrium values for the endogenous variables. Thus, each endogenous variable is a function of all the exogenous variables in each district. See pp798-99.

children to public or private schools, and (iii) how to vote on the local property tax rate. The property tax rate (t_j) in each community is determined by majority rule and the state income tax rate (τ_j) is set to balance a state-level budget constraint. Finally, the household's expenditures must satisfy their individual budget constraint:

$$(21) \quad (1 - \tau_j)y_i + w_i = b_{i,j} + (1 + t_j)p_j + T.$$

A household's disposable income is equal to their after tax income plus their non-taxable income, w_i . This is spent on consumption of the numeraire, housing, and possibly tuition for a private school, T .

Ferreya estimates all of the parameters of her model using a combination of school district data and Census data from the metropolitan areas of New York, Chicago, Philadelphia, Detroit, Boston, St. Louis, and Pittsburgh. A minimum distance estimator is used to match predicted and observed levels of several variables describing these metropolitan areas and the school districts located within them: (i) average household income, (ii) average housing price, (iii) average expenditures per student in public schools, (iv) the share of households with children in public schools, and (v) the share of households with children in Catholic schools.

4.4 *Strengths and Weaknesses of Empirical Sorting Models*

The *PC*, *RU*, and *CS* frameworks differ in how they: (a) define the set of choices faced by each household; (b) specify the shape of the preference function; and (c) develop instruments to control for endogenous amenities. These modeling decisions are important because they place conditions on the estimates of parameters describing preferences for amenities. Each implies a set of restrictions that combine with observed behavior to determine what can be learned from the model. Estimates for the distribution of structural preference parameters will, in turn, shape predictions for how households, market prices, and endogenous amenities would react to a prospective policy shock. Table 1 summarizes the key modeling decisions that differentiate these three models from each other and from conventional hedonic approaches to demand estimation.

4.4.1 *The Choice Set*

In a discrete choice framework, estimates for the structural preference parameters can be sensitive to the definition used for the choice set and to the composition of choices within that set (Kenneth E. Train 2003; Kuminoff 2009). An advantage of the RU and CS models is that they facilitate describing the choice of a house as being related to choices in other markets. In the CS model, households may be forced to choose between inexpensive housing in communities with weak public schools and expensive housing in communities with strong public schools. Faced with this tradeoff, some households may choose to purchase a large house in a less expensive community and then pay to send their children to a private school. Similarly, the RU framework recognizes that a household's choice of a home may depend on where its primary earner works. Treating each worker's job location as fixed, this information can be used to calculate the commuting distance to each potential house location. Thus, every working household in their model faces a unique choice set.²⁸

Bayer, McMillan, and Reuben's (2004) application of the RU model is the only one to acknowledge the connection between housing and labor market choices. However, their random utility model lacks a budget constraint. Elsewhere in the discrete choice literature, it is common to assume that every alternative in the choice set is also in every consumer's budget set. This assumption seems plausible in markets for breakfast cereal (Aviv Nevo 2001), laundry detergent (Igal Hendel and Nevo 2006) and even personal computers (Bajari and Benkard 2005). It is not plausible in the market for housing.²⁹ Few households are sufficiently wealthy to be able to afford every home in a metropolitan area. The PC and CS models are consistent with this observation. In the PC framework, for example, the identifying assumption is that each household is able to purchase at least some quantity of housing in every community. This assumption seems plausible for communities that contain a variety of differently sized houses sold at a wide range of prices. The unified school districts that have served as communities in most applications of the PC framework satisfy this requirement.

There are at least two ways to incorporate a budget constraint into the RU framework. Adapting the logic of the PC framework, Bayer, Keohane, and Timmins (2009) replace p_n in the

²⁸ Bayer and Timmins (2007) demonstrate that variation in the choice set also increases the power of instruments based on substitute attributes.

²⁹ Banzhaf and Smith (2007) explore the implications of alternative assumptions for how "affordability" conditions the choice set in their estimation of bid functions.

indirect utility function with an index of community-specific housing prices and assume households are free to choose continuous levels of structural characteristics in every community, allowing h_n to be optimized out of the indirect utility function. Alternatively, Takeuchi, Cropper, and Bento (2008) replace p_n with $\ln(y_i - p_n)$ in order to set the probability of occupying a house equal to zero if its annualized price exceeds the household's income. The challenge with implementing this approach is measuring the dimensions of a household's wealth that constrain its purchase of a house. Ideally, this measure would reflect past savings, current income, and expected future earnings.

4.4.2 Preference Heterogeneity

To compare the depiction of preference heterogeneity in the three models consider an example where communities differ in two observable amenities, air quality and school quality. In all three models the utility function contains a separable, linear sub-function of amenities. Let γ represent the relative preferences for these amenities. Using this notation, the three sub-functions can be represented as follows:

$$(22) \quad \begin{array}{l} (PC) \quad \gamma_1 \cdot school_j + \gamma_2 \cdot air_j + \bar{\xi}_j \\ (CS) \quad \gamma_{i,1} \cdot school_j + \gamma_2 \cdot (\gamma_3 \cdot air_j + \bar{\xi}_j) + \varepsilon_{i,j} \cdot \\ (RU) \quad \gamma_{i,1} \cdot school_j + \gamma_{i,2} \cdot air_j + \bar{\xi}_j + \varepsilon_{i,j} \end{array}$$

$\bar{\xi}_j$ measures the mean effect of unobserved amenities at location j , while $\varepsilon_{i,j}$ is an idiosyncratic deviation from that mean.

In the *PC* specification, every household is required to have the same relative preferences for every amenity, observed and unobserved. Everybody trades air quality and school quality at the same rate. The *CS* specification relaxes this depiction of preference heterogeneity by allowing households to differ in their relative preferences for school quality. However, households still have the same relative preferences for air quality and $\bar{\xi}_j$. The *RU* specification provides the most general depiction of preference heterogeneity.

Preference heterogeneity determines the way in which households perceive communities to be different. When households share the same relative preferences for each pair of amenities in the *PC* specification, every household must also agree on a common ranking of communities

by the overall quality of amenities they provide. The notion that a set of differentiated communities can be unanimously ranked by quality is analogous to Lancaster’s (1979) description of “vertical” product differentiation. In contrast, when households differ in their relative preferences for amenities, as in *CS* and *RU*, they may also differ in the way they rank communities—a situation analogous to Lancaster’s description of “horizontal” product differentiation. The distinction between vertical and horizontal differentiation is important because the two concepts differ in their implications for substitution patterns (Simon P. Anderson, André de Palma, and Jaques-François Thisse 1992). When households agree on the ranking of communities by overall amenity provision in the vertically differentiated case, they must also agree on the opportunities for spatial substitution. For example, if asked to identify the two closest substitutes for community j , every household in the *PC* specification will select the two adjacent communities in the ranking. Horizontal differentiation allows more diversity in substitution possibilities. Since households in the *CS* and *RU* specifications can differ in how they rank communities, they may disagree on which communities represent the closest substitutes for j . The idiosyncratic elements in preferences guarantee that each location is a substitute for every other location.

PC, *CS*, and *RU* also differ in how they restrict substitution between amenities and private goods. All three models assume that different amenities are perfectly substitutable. This conclusion follows from the additive form of the indices in (22). The CES specification used by *PC* allows the relationship between amenities and private goods to in principle range from perfect substitutes to perfect complements. In contrast, the linear utility functions in *RU* and *CS* restrict amenities to be perfect substitutes for the private good numeraire. This restriction is quite common in random utility models of the choice among differentiated goods. However, it is not a requirement for consistent estimation. Specifications relaxing the assumption of perfect substitutability would be desirable in future research.

Finally, the idiosyncratic elements in preferences in the *CS* and *RU* specifications deserve some comment. It seems plausible that households will differ in their preferences for individual locations in ways the econometrician cannot observe. A household may, for example, be attached to a particular community because its family and friends live there. Failing to account for these effects may cause the *PC* specification to be biased. However, the extra flexibility provided by adding the J idiosyncratic elements, $\varepsilon_{i,1}, \dots, \varepsilon_{i,J}$, comes at a considerable cost because the

model is only identified by the maintained assumption that these idiosyncratic shocks are drawn from a series of independent, identically distributed type I extreme value distributions. The composite of these different restrictions suggests that the vertical/horizontal modeling choice can be viewed as offering a bias/variance tradeoff. For example, suppose that horizontal differentiation is the “true” form of preference heterogeneity. By restricting relative preferences, vertical differentiation would bias the *PC* estimator for preferences, as well as the conclusions that would be drawn about demand curves and welfare measures. Horizontal differentiation eliminates the restriction that causes bias, but the added dimensionality of preferences increases the scope for untested distributional assumptions to drive the estimates.³⁰

4.4.3 *Instruments*

PC, *RU*, and *CS* each suggest a different IV strategy. *CS* directly models the production process for the endogenous amenity, *PC* develops instruments based on functions of the community income rank, and *RU* develops instruments based on functions of the exogenous attributes of substitute locations.

The “income rank” and “substitute attribute” instruments are closely related. They both rely on a similar assumption about the role of unobserved amenities in household location choice. Recall that the validity of the income rank instruments rests on the assumption that unobserved amenities do not affect the ranking of communities by income.³¹ This assumption is most likely to hold when unobserved amenities play a minor role in determining where people choose to live. As the importance of unobserved amenities increases, the strength of the substitute attribute instruments decreases. This is because the power of the substitute attribute instruments depends on the extent to which observable exogenous amenities drive household location choice. Thus, the relevance of the substitute attribute instruments used by *RU* is linked to the validity of the income rank instruments used by *PC*.

The advantage of the production function approach to addressing endogenous amenities

³⁰ There is a similar bias/variance tradeoff between a representative agent framework and the *PC* model, in which the distribution of income and preferences is identified by the assumption that they are jointly drawn from a bivariate normal distribution.

³¹ The “validity” of an instrument is determined by the validity of the assumptions used to establish its relationship to the econometric error term. An “invalid” instrument is one that is correlated with the error and, therefore, does not support consistency of the estimator. An “irrelevant” instrument is uncorrelated with the instrumented variable. A weak instrument has low relevance.

in the CS model, illustrated by Ferreyra (2007) and Calabrese et al. (2006), is that it does not require any assumption about the relative importance of unobserved amenities. One avoids the need to construct an instrumental variable by modeling the mechanisms describing how a positive draw on \bar{g}_j or $\varepsilon_{i,j}$ raises the level of school quality in community j or the price of housing. This might seem to be an ideal basis for general equilibrium simulation of policies that affect school quality. However, it presents a tradeoff. The production function approach accepts as a maintained condition a specific form for the amenity's production function and uses a full solution method; i.e. it computes an equilibrium on every iteration of the estimation routine. This approach ties the consistency of the estimator to the specification for the amenity production function, whereas the instrumental variables approaches used by *PC* and *RU* are consistent as long as the instruments are valid. Furthermore, *CS* implicitly assumes that unobserved amenities do not vary systematically across school districts. To the extent that unobserved amenities are spatially correlated across districts, they will be absorbed by the district fixed effects during the first-stage estimation of \bar{g}_j . This will conflate the estimated effects of school quality and unobserved amenities.

Finally, it is important to note that instrumental variables need not arise from the structure of the equilibrium. In principle, any source of exogenous variation in an endogenous amenity can serve as an instrument, just as in reduced form hedonic methods. It can be difficult to find an exogenous source of variation in an amenity. Moreover, quasi-experimental designs that exploit this variation can have other drawbacks. This is illustrated by Bayer, McMillan, and Rueben (2004) and Bayer, Ferreira, and McMillan (2007), both of which use Black's (1999) boundary discontinuity design to isolate exogenous variation in school quality in the *RU* framework. While the boundary discontinuity approach has intuitive appeal as an identification strategy, it involves discarding the data on all houses located more than a short distance from a boundary.³² Often this includes the vast majority of houses in the study region.

4.5 Calibration as an Alternative to Estimation

³² In section 6.3.2 of their review of regression discontinuity designs, David S. Lee and Thomas Lemieux (2010) discuss threats to the internal validity of the boundary discontinuity strategy. They suggest several ways in which urban development and household sorting may reduce the comparability of houses on opposite sides of a boundary. The same mechanisms may compromise external validity by reducing the comparability of households who chose to locate inside and outside of the boundary zones.

Equilibrium sorting models require specific assumptions about preference and production functions and, sometimes, the prior specification of numerical values for key parameters. If our objective is to characterize distributions of outcomes that result from a particular policy intervention and depend on features of that structure, then the existing data may be insufficient in some applications to estimate every relevant parameter. The sorting literature contains several examples where analysts have used mixtures of parameter estimates from the literature to provide a partial description. In these cases, to close the model, values for the remaining parameters are selected through an estimation-like process to match the observed and simulated distributions of variables generated by the model. The calibrated model can support prospective policy analysis or be used to evaluate past policies and conduct model validation. Table 1 provides some limited examples of calibration studies. We use papers by Fernandez and Rogerson (1998) and Epple and Ferreyra (2008) to illustrate the differences between calibration and estimation.

Epple and Ferreyra's (2008) formulate a stylized model of the urban landscape, prove the existence of sorting equilibria before and after a specific policy reform, and then consider a change in financing of public education and a change in expenditures, holding per pupil spending constant. Six analytical predictions are derived. Predictions for the ordering of prices (and measures for education), stratification of income, and demographic composition are evaluated based on what can be observed when the reform (in a more detailed format) was actually implemented. They are able to formulate testable predictions about policy outcomes without having to first choose (or estimate) specific parameter values for parametric functions. Their analysis evaluates the importance of general equilibrium effects for a reform in public school financing and tax policies in Michigan. The reform lowered property tax burdens and reallocated revenue to school districts. It increased revenue for low revenue districts and capped revenues of high revenue districts. Their findings confirm the importance of general equilibrium effects, especially when they reflect preference heterogeneity and the diversity in household adjustments.

Fernandez and Rogerson's (1998) approach to calibration is closer to estimation. They develop a two-community, two-period, overlapping generations model where the "old" cohort makes choices about the quality of education for the "young" cohort in period one, which contributes to the young cohort's income when they grow old in the second period. Each indi-

vidual is restricted to make a single, permanent, location decision. The collective choices made by the current generation influence the next generation's distribution of income. This allows the effects of current policies on both current and future income distributions to be considered.

The interaction between households is described as a three stage game. In the first stage, the first generation agents simultaneously choose one of the communities. In the second stage, given a community decision, property tax rates are selected based on majority rule. Agents know community composition when voting. Aggregate tax receipts determine education quality. In the third stage, individuals make housing and consumption choices. The model can be solved with backward induction focusing on a boundary indifference condition that defines the fraction of the population in each community.³³ A single crossing condition assures stratification and allows them to characterize the equilibrium.

Their calibration procedures involve selecting parameter values for the preference, education production, and housing supply functions, and for the distribution for the idiosyncratic effects that are implied by the model's assumptions using the steady state solution values for these parameters (see Table 1 in Fernandez and Rogerson 1998). The fitting criterion appears to be an equally weighted distance function. While this approach is similar to a minimum distance estimator, there is only one "observed" value for each parameter, a best estimate from the literature. A key feature of this calibration strategy is that outcomes that were not used during the calibration process were later used to judge the fit of the calibrated model. The Fernandez-Rogerson (1998) and Epple-Ferreira (2008) examples illustrate how the process of model calibration can advance our understanding of equilibrium sorting in ways that complement structural estimation and contribute to policy evaluations when complete data for estimation of these types of models are not available.

4.6 *Tradeoffs in Implementation of Sorting and Reduced Form Models*

Equilibrium sorting theory suggests some broad guidelines for research design. For example, maintaining the assumption of free mobility would imply that the geographic extent of the market should be limited. Likewise, the individual outcomes described by models of sorting

³³ Associated with this fraction is a measure of utility associated with the lowest income person in the high-income community and the highest income person in the low-income community.

behavior and hedonic equilibria can be measured more directly with micro data than with aggregate data. With a static definition for the equilibrium that holds individual income and preferences fixed, one would prefer to have a fixed number of cross-sectional observations drawn from a shorter temporal interval. Based on this logic, many hedonic and sorting applications define “the housing market” to be a single metropolitan area observed over a few years where the spatial unit of observation, a housing sale, matches someone’s actual location decision. This definition may provide a reasonably good match to the theory but, unfortunately, there may be a problem. The problem arises when the definition of a market is incompatible with the analyst’s preferred instrument for an endogenous amenity. Variation in the instrument may only be observable over a larger spatial area, a longer time period, or a lower level of spatial resolution. Thus, to use the instrument, the analyst needs to manipulate, at least implicitly, the definition for “the market” to strengthen the assumptions needed to interpret parameter estimates in terms of the underlying sorting model.³⁴

The tradeoff is especially apparent in recent hedonic models that have adapted the logic of program evaluation to housing markets. Analysts have identified credible sources of exogenous variation in amenities over space (Black 1999) and time (Davis 2004, Chay and Greenstone 2005, Greenstone and Justin Gallagher 2008). In order to use this variation to draw inferences about MWTP from hedonic price functions describing sorting equilibria, the applications must assume that the shape of the price function is unaffected by large changes in the amenity of interest (Davis 2004). In other applications one needs to add the further assumption that the free mobility and perfect information assumptions of the hedonic model apply to the contiguous U.S. (Chay and Greenstone 2005). In some applications of the program evaluation logic, the research design can be implemented using micro data (Black 1999, Davis 2004) but in other cases conducting a national analysis or using the preferred instrument requires summary statistics such as median or mean values for houses in a Census tract to be used to represent equilibrium prices (Chay and Greenstone 2005, Greenstone and Gallagher 2008). This arises because the instru-

³⁴ All else constant, a natural experiment or a quasi-experiment would arguably provide the most credible research design for answering any specific question about sorting in housing markets. Imbens (2010) states that his main concern with this line of reasoning is that there are many important questions—such as those involving general equilibrium effects—that cannot be answered directly with simple experiments. In these situations, Heckman (2010) warns that the parameters that can be identified may conflate the economic measure of interest with other contaminating influences. See both papers for an insightful discussion of research design cast in terms of the broader econometric literature on program evaluation.

ment varies at spatial scale that requires a broad geographic area be used in defining the extent of the market. Do these assumptions alter the appropriate economic interpretation of the identified parameters? It is important to consider this question because some structural models also use aggregate data (e.g. Ferreyra 2007) and insights from the program evaluation literature are increasingly being used to modify the research designs used to estimate structural preference parameters (e.g. Bayer, Ferreira and McMillan 2007, Bayer, Keohane, and Timmins 2009, Galiani, Murphy and Pantano 2012).

In general we should begin by recognizing that all reduced form measures of the effects of a housing market treatment that is changed over time recover measures for the “capitalization effects” of that change. Kuminoff and Pope (2012) demonstrate that, in some cases, capitalization effects will have a welfare interpretation comparable to that of Rosen’s (1974) hedonic model. In other cases it is unclear how to interpret these effects due to the economic implications of the assumptions maintained. For example, invoking Rosen’s welfare interpretation relies on the assumption that the gradient of the hedonic price function is constant over time (i.e. there is no temporal dimension to his model). Intuitively, one might expect the hedonic gradient to be approximately invariant to changes in amenities that are “small” or “localized” in the sense that they occur over a small portion of the market (Palmquist 1992). On the other hand, small changes in one amenity may trigger tipping effects that produce large changes in other features of equilibria (Bénabou 1996b, Sethi and Somathathan 2004). If the migration patterns that follow a small change in environmental quality nudge the minority share of a neighborhood beyond its tipping point, for example, then there may be large subsequent changes in racial composition and other endogenous amenities that alter the hedonic gradient (David Card, Alexandre Mas, and Rothstein 2008, Banzhaf and Walsh 2010). Understanding the transitional dynamics between equilibria is an important area for research.

Regardless of whether or not they have a direct welfare interpretation, accurate measures for amenity capitalization effects have inherent value. These effects matter to homeowners, to renters, and to the beneficiaries of programs funded by property tax revenue. Under well-defined conditions, reduced form methods can estimate average capitalization effects and local average capitalization effects (for a general characterization of ATE and LATE, see Imbens and Wooldridge 2009). However, if the analyst’s objective is to use capitalization effects to assess

the welfare implications of a past event, then at least three questions must be addressed. Do the data describe a single geographic market connected by a common hedonic price function? Was the gradient of the price function constant over the duration of the study period? Are the “treated” houses in the sample representative of the population of interest? If the answer to any one of these questions is ‘no’, then the ideal measures of capitalization effects cannot be interpreted as parameters describing a hedonic price function or used for welfare measurement.³⁵ Knowledge of sorting behavior can help to answer these questions. For example, Gamper-Rabindran and Timmins (2011) and Kuminoff and Pope (2012) demonstrate that assumptions about the spatial extent of the market and the temporal stability of the hedonic gradient can (and should) be tested within the context of quasi-experimental research designs.

5. Using Sorting Models for Prospective Policy Evaluation

Most policy evaluations translate a proposed change in product quality, product quantity, or market institutions into an equivalent price change for an average producer or consumer, holding fixed all other features of the model. Arnold C. Harberger’s (1964) approximations can be described in these terms. He translates tax changes into equivalent price changes and then into measures of excess burden. Similarly, the conventional strategies used to estimate the value of new goods (or to incorporate them into price indexes) treat new commodities as if they have always been part of the choice set, but that prior to their “introduction”, they were simply above everyone’s choke price.³⁶ In the context of non-market goods, Smith and Banzhaf (2007) suggest that the concept of weak complementarity can be interpreted as allowing the definition of price equivalents for quality changes.³⁷ In general, exogenous changes in non-price attributes of a product can be translated into price equivalents, given assumptions about preferences or technology.

When we consider markets where heterogeneous consumers sort over differentiated goods with endogenous characteristics it becomes more difficult to define price equivalents for policy changes. Moreover, the concept of price equivalents may be insufficient to quantify the

³⁵ Klaiber and Smith (2012) demonstrate how to use a calibrated model of sorting behavior to explore the ability of a quasi-experiment to define treatment and control groups that serve to identify the underlying economic parameters of interest.

³⁶ See E. Rothbarth (1941) and Walter E. Diewert (1993) for historical discussion and Hausman (1997) for an example.

³⁷ They are equivalent (for an individual) in terms of Hicksian welfare measures to changes in the non-market good.

outcomes that matter to households and policymakers.

Consider a prospective policy that aims to change one amenity from its current level, g_j , to some new level g_j^* . The size of the change may vary across space. This is certainly true for national policies that define minimum standards for amenities, such as the No Child Left Behind Act, the Clean Air Act, and the “Superfund” program for cleanup of hazardous waste sites. It may also apply to local policies. For example, a local property tax assessment to fund the preservation of open space may affect households differently depending on where they live relative to preserved parcels.

There are several questions one may want to ask about the possible outcomes of such a policy. First, how will it affect market prices in targeted locations and non-targeted locations? Second, will the policy induce some households to move and, if so, how will the resulting migration patterns affect the levels of endogenous amenities at each location? Third, how much would households in each location be willing to pay for these outcomes? Finally, how are the gains and losses from the policy distributed across households? Are they disproportionately borne by specific demographic groups? All of these questions are equally relevant if we instead consider a policy that alters the dimensions of the choice set, or a policy that alters market institutions such as the existing schedule of taxes and subsidies.

This section discusses how sorting models change the way the evaluation task could be undertaken relative to the conventional hedonic literature. Parts 1 and 2 explain in general terms how these models can (and have) been used. Parts 3 and 4 cover three leading examples from the empirical literature (education, air quality, and land use) where moving from hedonic analysis to sorting models has expanded the scope of policy evaluations.³⁸ We conclude with our evaluation of the types of questions that the current generation of sorting models is capable of addressing.

5.1 *Hedonic Analysis of Public Policy*

Sherwin Rosen’s 1974 paper currently has over 6000 citations (based on Google’s citation gadget 12/21/2012). Empirical studies using his model to evaluate policy are typically

³⁸ The authors of most of the sorting studies that we cover first estimate or calibrate a structural model before using it to simulate counterfactual equilibria. Our coverage in this section focuses on the results of the simulation exercises.

limited to the first stage of his procedure—estimating the hedonic price function.³⁹ As we noted earlier, given Rosen’s assumptions the price function reveals each household’s marginal willingness to pay for an amenity, as defined in (13).

Rather than report the entire distribution of MWTP, the analyst typically emphasizes a single statistic, such as the mean or median. Mean or median MWTP is then used to construct a “back of the envelope” approximation to a statistic with some policy relevance. For example, Linden and Rockoff (2008) use hedonic MWTP to estimate the “victimization cost” of a sexual offense, Davis (2004) estimates the statistical value of a case of pediatric leukemia, Chay and Greenstone (2005) estimate the willingness to pay for large reductions in particulate matter, and Greenstone and Gallagher (2008) test whether the Superfund program is “worth it” based on the average cost for the cleanup of Superfund sites and the estimates from their model.

There are at least four important limitations to using approximations to hedonic price functions to evaluate policy. First, the assumptions allowing the hedonic gradient to be translated into MWTP are quite strong. In particular, households must be free to select an amenity bundle from a continuous joint distribution.⁴⁰ The direction of the bias from violating this assumption is indeterminate. Second, measures of MWTP cannot be used to calculate measures of compensating or equivalent surplus unless (a) demand curves are perfectly elastic over the range of the change, or (b) all households are identical. Third, these calculations assume the ways households adjust their behavior in response to the policy are offsetting so the hedonic price function does not change. Finally, hedonic approximations do not provide a consistent basis for predicting how a new policy (that is different from existing policies) will affect market outcomes. For a sufficiently large shock to an amenity, households may respond by changing their location, which will affect the hedonic equilibrium, community demographics, other endogenous amenities, and feed-back into welfare measures. For example, Banzhaf and Walsh (2008) find that when industrial facilities that emit toxic chemicals move into a neighborhood, some residents move out. When these facilities move out, households with different demographic characteristics move in. The price adjustments needed to clear the housing market following these shocks will affect the welfare of homeowners and renters. These outcomes cannot be

³⁹ See Palmquist (2005b) for a review of the empirical hedonic literature.

⁴⁰ Many of the structural estimators described in section 4 also assume the existence of a hedonic price function, but have no need to invoke the continuity assumption in order to identify parameters describing preferences.

predicted from the hedonic price function.⁴¹

Given these limitations, it is natural to ask whether hedonic MWTP might be used to construct an upper or lower bound on the benefits of a policy. Unfortunately, the answer, in general, is no. Suzanne Scotchmer (1985) demonstrated that the information contained in marginal implicit prices is insufficient to predict how markets will adjust to a future change. Bartik (1988) and Yoshitsugu Kanemoto (1988) investigated the possibility of using the information in marginal implicit prices to calculate ex ante bounds on ex post welfare measures. While they both report positive results, their conclusions are contradictory. The reason is that they differ in how they define the initial equilibrium and the possibilities for adjustment. Under a restrictive set of conditions, Kanemoto proves that ex ante welfare measures will overstate the benefits from an amenity improvement. The requirements for his proof include homogeneous households, only two types of houses, and that the improvement is funded from tax revenue. These restrictions are relaxed in Bartik's model. He argues that ex ante welfare measures will understate the benefits from an improvement when heterogeneous households face a diverse set of housing opportunities. This contrast highlights how the idiosyncratic features of an application may determine the direction of the difference between ex ante and ex post welfare measures.⁴² It is also worth noting that both studies assume non-targeted amenities are strictly exogenous. With endogenous amenities it seems likely that "anything goes".

Very few empirical studies have implemented Rosen's second stage to estimate the demand for an amenity. The barrier to reduced-form estimation has been the need for data from multiple markets on housing transactions, household demographics, and instruments.⁴³ The traditional concern with structural estimation is that restrictions on preferences are arbitrary. Recent microeconomic innovations by Ekeland, Heckman, and Nesheim (2004), Heckman,

⁴¹ It is tempting to think the hedonic price function could (at the very least) be used to predict the price that would be charged for a new choice alternative. However, the price function simply describes the current market equilibrium. Introducing a new choice alternative may change the entire price schedule. Bajari and Benkard (2005) discuss this same point in the context of markets for differentiated private goods.

⁴² Heckman, Matzkin and Nesheim (2010) reinforce this point in distinguishing the restrictions for point identification of parameters from those required to measure welfare effects. They repeat the Kanemoto conclusion for a case of quasilinear preferences that admits a diverse group of alternative specifications (see their eq 4.1). They consider how a change in the hedonic price function for specific households who experience the change in the amenity directly would influence conclusions as compared to using the price function to evaluate welfare for those who might sort in response to a change in the amenity. The later are included in Bartik's assessment.

⁴³ Of course, one could also use parametric restriction as with Mendelsohn (1985) and more recently Sudip Chattopadhyay (1999).

Matzkin, and Nesheim (2010), and Bajari and Benkard (2005) have reduced the barriers to second-stage estimation. The first two papers focus on non-parametric strategies. They provide preliminary evidence that data from multiple markets may be unnecessary to implement the reduced-form approach. Bajari and Kahn (2005) relax some of the rigidity of the structural approach by allowing individual households to differ in their tastes for each characteristic. Bajari and Benkard (2005) provide a more general characterization of this logic that also relaxes the need for continuous choice sets and perfect competition.⁴⁴

5.2 Using Equilibrium Sorting Models for Policy Evaluation

A structural model that provides estimates of household preferences allows the Hicksian compensating surplus to be measured for a specified change in amenities. Equation (23) defines this partial equilibrium measure of willingness to pay (WTP_{PE}) for a prospective change in the targeted amenity.

$$(23) \quad V(g_{1,j}, g_{1,j}, p_j; \alpha_i, y_i) = V(g_{1,j}^*, g_{-1,j}, p_j; \alpha_i, y_i - WTP_{PE}).$$

Baseline and new levels of the targeted amenity in location j are denoted by $g_{1,j}$ and $g_{1,j}^*$. This is a partial equilibrium measure of the change in welfare in the sense that prices, income, and the levels of other amenities ($g_{-1,j}$) are held fixed. The distribution of WTP_{PE} could be calculated from estimates for the joint distribution of income and preferences derived from either a structural hedonic or an equilibrium sorting model.

Sorting models extend the scope for evaluation by allowing the analyst to predict how markets might adjust to the change. In the context of local public goods, the models allow relocation and price adjustment as the (assumed) supply of housing and the demand implied by sorting are equalized in each location. Moreover, depending on how the model represents amenities, their supply may also be treated as an endogenous outcome stemming from the equilibrium adjustments. In the simplest case this connection can be “technical”. For example, the amenities associated with open space may be expressed as a function of the share of protected and vacant land in the neighborhood (Walsh 2007) or the production of school quality may be

⁴⁴ With few exceptions, estimates for demand curves and welfare measures based on structural and reduced-form models continue to maintain the assumptions of *perfect information*, *free mobility*, and *no discrimination*. Relaxing these assumptions, or investigating how they affect the interpretation of results, is an interesting topic for research.

related to student teacher ratios that adjust as households with school-age children move (Klaiber and Smith 2012). Alternatively, expenditures of local public goods may be determined by voting (Ferreyra 2007). In this case migration may alter the median voters in some communities, changing voting outcomes, and so forth.

A “general equilibrium” measure of individual willingness to pay consistent with the preceding examples can be defined as

$$(24) \quad V(g_{1,j}, g_{1,j}, p_j; \alpha_i, y_i) = V(g_{1,k}^*, g_{-1,k}^*, p_k^*; \alpha_i, y_i - WTP_{GE}).$$

A key distinction from (23) is that willingness to pay is now evaluated using the household’s new location, k , which may or may not be the same as its initial location. Following convention we label this welfare measure as “general equilibrium” in recognition of its allowance for market adjustment.⁴⁵ However, the label “multiple market” might be more apt. The measure of WTP_{GE} in (24) recognizes the interconnectedness between the supply and demand for public and private goods in each of the J communities. However unlike computable general equilibrium models we are holding incomes fixed, as well as the price of the composite numeraire.

There are many interesting aspects of how sorting models are actually used to define the new equilibrium that might follow a policy change. We have selected three classes of issues to discuss: (a) the information used to close the model; (b) the assumptions about sources of “friction” in the market; and (c) the potential for multiple equilibria.

5.2.1 *Closing the Model*

Solving for a new equilibrium typically requires more information than was used for the estimation. Recall that empirical models are estimated using data for intervals over which the number of households, the number of houses, and the amounts of each amenity are assumed to be fixed. To solve for a new equilibrium, the analyst must consider how the supply of each might change.

Given some additional supply information, numerical methods are used to solve for equi-

⁴⁵ Equation (24) is most directly related to the pure characteristics sorting models. The formulation would be a bit different for the random utility models. The reason is that its equilibrium condition does not predict the selection of a specific location. Instead it provides probabilities that each choice location will be selected. Equilibrium is defined by a vector of prices at which the expected housing demand in each location equals supply. This logic translates into measures of expected willingness to pay. In our view this is not an important variation on the central logic of the sorting model. The use of expectations smoothes the computation of equilibria and welfare measures, similarly to Berry, Levinsohn and Pakes (1995).

librium housing prices, amenities, and location choices that simultaneously satisfy the following three conditions.⁴⁶

$$(25.a) \quad V(g_k^* p_k^*, y_i, \alpha_i) \geq V(g_m^*, p_m^*, y_i, \alpha_i) \quad \forall \quad i, m \neq k .$$

$$(25.b) \quad H_k^S(p_k^*) = H_k^D(p_k^*) \quad \forall \quad k .$$

$$(25.c) \quad g_k^* = f[g_k, g_{-k}^*, p_k^*, p_{-k}^*, R(\alpha, y, d)] \quad \forall \quad k .$$

The first condition simply states that each household must select the location that maximizes its utility. Condition (25.b) requires the supply of housing to equal demand in each location. In the pure characteristics model, for example, housing demand is calculated by aggregating over the individual demand curves in (17). Supply is treated differently depending on the application. Smith et al. (2004) restrict supply to be perfectly inelastic, Sieg et al. (2004) calibrate the supply curve using a range of elasticities, and Walsh (2007) estimates a land supply model using the historical records of residential development in his application. Finally, condition (25.c) expresses the new levels of amenities in community k as a function of their baseline levels, the new levels of amenities in all other communities, g_{-k}^* , housing prices, and $R(\alpha, y, d)$, the joint distribution of income, preferences, and demographic characteristics describing the population of households. For example, the production of school quality in Ferreyra's (2007) model depends on social interactions and voting outcomes, both of which depend on the location choices made by households which, in turn, depend on all of the features of the spatial landscape.

The population of households in the model is usually treated as fixed. In overlapping generation models such as Fernandez and Rogerson (1998), the distribution of income may differ from generation to generation, but the size of the population itself does not change. In other words, there is assumed to be no immigration or emigration from the study region. This assumption influences predictions for capitalization. Several studies have reported simulation results where housing prices decrease in communities that experience improvements (e.g. Sieg et al. 2004). Intuitively, because the model is closed, a policy that improves amenities in every location can still make some locations *relatively* less attractive.

⁴⁶ The details of numerical solution procedures vary with the model and application. For the pure characteristics model see Sieg et al. (2004), Walsh (2007), and Kuminoff and Abdul S. Jarrah (2010). For random utility models see Bayer and Timmins (2005), Timmins (2007), and Klaiber and Phaneuf (2010b). See Ferreyra (2007) for details on her general equilibrium model.

Finally, an assumption is required about who collects the capital gains (or losses) from housing transactions. Applications in the RU and PC traditions have treated households as renters, assuming that changes in property values accrue to absentee landlords.⁴⁷ In contrast, CS models treat all households as homeowners. For example, Nechyba (2000) endows each household with a specific house in the baseline equilibrium. When a homeowner relocates in response to a policy change he sells his house, collects the capital gains, and purchases a different house in his new location. Thus, the distinction between owners and renters may be important for general equilibrium welfare calculations. All else equal, homeowners cannot be made worse off from quality improvements. However, a renter with the same preferences and income may be worse off if their rent increases by more than their willingness-to-pay for the improvement.

Census micro data differentiate between renters and owners, and the current structural estimators are capable of recognizing that owners and renters may differ in their preferences. However, the current general equilibrium models do not allow for transitions between renting and owning. This is a difficult and important challenge for future research. Modeling the decision of whether to rent or own would likely require data on assets and access to credit, as well as a dynamic model that includes moving costs and forward looking behavior.⁴⁸ We return to this task in our discussion of research frontiers in section 6.1.

5.2.2 *Potential Sources of Friction*

The first generation of “general equilibrium” applications has continued to maintain the free mobility assumption that is embedded throughout the empirical hedonic and sorting literatures. There are no explicit physical or wage-related costs of moving to a new location. Furthermore, the applications to date have abstracted from transitional dynamics. Numerical simulations are implemented as if markets adjust instantaneously. One could argue this approach is consistent with interpreting predictions for market outcomes as features of long run equilibria. On the other hand, the models do not suggest a procedure for discounting. It is certainly clear

⁴⁷ An exception is Daniel G. Hallstrom and Smith (2003) who develop a pure characteristics model where all gains or losses accrue to the occupants of those houses in the baseline equilibrium. The prospective gains or losses influence their decision for whether to relocate.

⁴⁸ The choice of whether to rent or own has equally important implications for reduced form models that use hedonic price functions as a basis for policy inferences. In principle, currently available data can easily be used to estimate different hedonic price functions for rental property and owner-occupied housing. However, welfare interpretations of the hedonic gradient based on Rosen (1974) abstract from credit constraints, moving costs, and forward looking behavior.

that the lack of friction in the current models has the potential to influence their findings. Several recent studies have underscored the need to address this issue.

Kuminoff (2009) used an analytical version of the pure characteristics model to illustrate how the free mobility assumption creates a false sense of precision in estimates for preference parameters and welfare measures. The direction and magnitude of biases will depend on baseline equilibrium conditions as well as latent distributions of preferences and moving costs. Bayer, Keohane, and Timmins (2009) were the first to propose a solution. They developed a discrete-choice analog to the first-stage estimation of a hedonic price function that controls for the average cost of moving between metropolitan areas. Controlling for moving costs had a dramatic impact on their estimates of MWTP for reductions in air pollution (as measured by particulate matter). Ferreira (2010) reports similar findings in an investigation of the implicit moving costs conveyed by property tax regulations in California. Finally, a recent study by Epple, Romano, and Sieg (2012) includes moving costs in an overlapping generation model with voting. After calibrating the model to Boston, they demonstrate that moving costs play an important role in the evolution of community demographics and tax revenue. Extending the current generation of models to include moving costs and transitional dynamics raises new issues that are just beginning to be explored. We discuss these issues in section 6.1.

5.2.3 *Multiple Equilibria*

Our earlier summary of theoretical properties of sorting equilibria focused on the types of conditions used to guarantee existence and uniqueness. Uniqueness proofs have relied on strong restrictions on the dimensionality of preference heterogeneity and the vector of endogenous outcomes. Frankly, the situations where equilibria are known to be unique are the least interesting for policy evaluation. Analysts have addressed multiplicity by conducting sensitivity analysis and proposing informal decision rules for choosing among equilibria.

It is difficult to guarantee that equilibria are unique in the empirical sorting models covered in section 5.⁴⁹ Of the three frameworks in that section, the pure characteristics model imposes the strongest restrictions on preference heterogeneity. It has been used in several policy

⁴⁹ As discussed earlier, Bayer and Timmins (2005) establish uniqueness for a specific RU model that restricts social interactions to take the form of congestion effects.

applications to solve for new equilibria (e.g. Smith et al. 2004; Walsh 2007; Klaiber and Smith 2012). None of these studies have reported multiple equilibria. It may be the case that the model's "vertical" ordering of households, combined with the market clearing condition for housing, is sufficient to ensure uniqueness. However this conjecture has not been proven. Nor has it been tested in a setting with multiple endogenous amenities.

Multiple equilibria have been found to arise when vertical differentiation is relaxed. Kuminoff and Jarrah (2010) recover several different equilibria from a relatively simple pure characteristics model where households have horizontally differentiated preferences over a vector of exogenous characteristics describing individual houses. Timmins (2007) notes the possibility of multiple equilibria for his random utility model of location choice in Brazil, which includes an agglomeration (or congestion) effect. Likewise Ferreyra (2007) confirms the possibility of multiple equilibria.⁵⁰

The possibility of multiple equilibria is relevant for policy evaluation, regardless of whether the analyst takes a partial or general equilibrium approach to analysis. Consider the first case. If one parameter vector can support multiple equilibria, the reverse may also be true. Multiple parameter vectors may each be equally capable of explaining the data that are used to estimate a model. Different parameter vectors may generate different predictions for benefit-cost ratios and other metrics used to quantify policy outcomes in a partial equilibrium setting.

Some models resolve the multiplicity issue by imposing enough discipline to point-identify the equilibrium reflected in the data. For example, Calabrese et al. (2006) prove that their model has a unique equilibrium, conditional on community populations and on the ranking of communities by mean income. Therefore, each parameter vector evaluated during their estimation routine has a single equilibrium that is consistent with the populations and income distributions observed in the data. Alternatively, if multiple parameter vectors are equally capable of explaining the data, then partial identification methods can be used to recover sets of parameter vectors, which can be translated into ranges for policy relevant metrics (Kuminoff 2010).

The literature has yet to arrive at a set of "best practices" for predicting how people and markets would adjust to a large change. Few studies have explicitly addressed multiplicity in the

⁵⁰ In particular, see footnote 10 in Timmins (2007) and footnote 13 in Ferreyra (2007).

general equilibrium context and those that have take different approaches to the problem. For example, Timmins (2007) solves for a new equilibrium using the iterative procedure developed by Bayer and Timmins (2005). Their procedure has a unique solution conditional on its starting value and will tend to converge to a solution that is closer to the starting value if multiple equilibria exist.⁵¹ By defining the baseline equilibrium as the starting value for his simulation, Timmins eliminates consideration of other potential equilibria that differ dramatically from the baseline. This approach is consistent with an assumption that markets have inertia. Kuminoff and Jarrah (2010) take a different approach. They simulate an unambiguous improvement in amenities, recover multiple equilibria, and then focus on the subset of equilibria that are consistent with the evidence on how housing markets adjust to changes in amenities. For example, they discard equilibria in which housing prices decline following an improvement to public school quality. Their main conclusions are robust to the choice among the consistent equilibria.

A more conservative approach would be to characterize all of the potential equilibria. If key policy outcomes are invariant to the choice among equilibria, there is no need to develop decision rules. This approach would be most applicable to situations where a strong prediction for the *direction* of change in a market outcome would be informative. It would also be consistent with developing a partial identification approach to using equilibrium sorting models for policy evaluation, following the logic of Charles F. Manski (2007).⁵² Finding ways to address the implications of multiple equilibria will become increasingly important as more flexible models are adopted. Recently, Bajari, Han Hong, John Krainer, and Denis Nekipelov (2011) developed an algorithm for recovering all of the equilibria in certain types of games considered in industrial organization. Their work could provide a starting point for the equilibrium sorting literature.

5.3. *Examples from the Empirical Literature*

Sorting models alter what might be termed the “landscape for policy evaluation” so that it more closely matches the ways decisions are made. A direct example of the effects can be seen

⁵¹ For example, consider two potential equilibria and suppose that one of the new equilibria lies in between the baseline and the other new equilibrium in attribute space. The Bayer-Timmins algorithm will converge to this intermediate equilibrium.

⁵² We return to a broader discussion of the prospects for using partially identified sorting models for policy evaluation in our coverage of research frontiers in section 6.5.

by comparing hedonic and sorting models for evaluating land use policy. Virginia McConnell and Margaret Walls (2005) report a detailed review of reduced form hedonic estimates for the price effects associated with proximity to “open space”. For models measuring the effects of open space using fixed effects (i.e. defining the impact based on whether a house is within a distance zone) the effects vary with type of open space from two percent of the house price (for urban parks) to sixteen percent (for natural areas). The same phenomenon could be represented with a continuous measure of distance but the results are generally smaller (when evaluated at mean distance). For contrast, consider the findings of a sorting model. Klaiber and Phaneuf (2010b) use a RU sorting model for the Twin Cities to consider increases in non-park open space in different locations. They demonstrate that a 2.5 percent increase in protected open space (with specific locations designated) in three different areas (inner city, urban fringe and outside the city), yields measures for willingness to pay that differ by up to 470% depending on whether we consider the effects of movement (and changes in property values) and who is impacted (all households versus those in the areas where land is designated as open space). Their counterfactual policy generates the increase in open space by converting privately owned agricultural and undeveloped parcels to publicly owned (non-park) open space. The stock of housing is held fixed. Of course new houses may be built in the longer run.

Walsh (2007) in a second example of a sorting model evaluating policies associated with open spaces demonstrates that efforts to preserve open space may have unintended consequences. In his PC sorting model, an increase in publicly owned open space may stimulate future urban development if an increase in the demand for housing in the improved areas raises the price of land by enough to induce the owners of vacant parcels to sell their land to developers.

Smith, Sieg, Banzhaf, and Walsh (2004) link a PC sorting model for Southern California to the air quality model developed for the Los Angeles Air basin by the EPA as part of an assessment of the expected implications of different compliance assumptions for 2000 and 2010. Ambient concentration measures are estimated on a 5km by 5km grid and are matched to the model’s study area. It is possible to compare the resulting spatially delineated benefit measures with the information for the region assembled by EPA. Indeed EPA’s (2010) most recent guidance document for conducting benefit cost analysis devotes sections of several chapters to

general equilibrium effects of environmental rules and identifies Smith et al. (2004) as the beginning of efforts to include benefits within a general equilibrium setting.

Finally, if we consider an educational policy, such as the decision to cut teachers in response to budget deficits, there can be an adjustment in educational quality that arises from sorting. If school quality is a function of teachers and class sizes, then the realized quality will depend on how households react to the policy. Klaiber and Smith (2012) answer this question, embedding an education production function in a simple model of sorting across school districts in Maricopa County, Arizona. They find that differential retention policies across districts results in declining educational quality for all but one of the districts. Parents move, class sizes increase at the schools without teacher cuts, and all school districts “share” in the losses that arise from teacher cuts in just a few of them.

5.4. *Are Sorting Models Ready for “Prime Time” Policy Evaluation?*

Sorting models have many potential uses in policy analysis. For example, Executive Order 12866 mandates the development of benefit-cost analyses for all major federal policies. Economic analysis can also help policymakers to understand the consequence of local governmental decisions, land use changes, and exogenous shocks to a region such as a plant closing, a hurricane, or a forest fire. In what follows we discuss three potential uses of sorting models: regulatory analysis, local government policy, and assessing the effects of shocks to a region.

5.4.1 *Regulatory Analysis*

While there have been only a few sorting models that have considered specific changes in federal regulations, the consistency in findings across models has been quite high, and where there are differences they seem to be readily explained. As noted in the case of air pollution there have been several applications. Two that can be directly compared are Sieg et al. (2004) (SSBW) and Tra (2010). They use the PC and RU sorting frameworks, respectively, to consider the benefits from the 1990 Clean Air Act Amendments for the Los Angeles area. Table 2 compares their results for MWTP as well as the partial and general equilibrium willingness to pay for the air quality improvements. Their findings are relatively consistent despite differences

in their regional scope, structure of preferences, and approach to measuring consumer surplus.⁵³

We also compute an approximate version of the Chay-Greenstone (2005) (CG) elasticity of housing expenditures to air pollution (designated as e here) by treating the general equilibrium WTP as the largest increase in annual housing expenditures a household would make for the improvement in air quality (particulate matter in the case of Chay and Greenstone and ozone for the other two studies). The CG estimates range from .20 to .35 in absolute magnitude (depending on model specification). Most of the results in SSBW match this range (with only the results for Ventura County larger) and Tra's estimates are at the high end of the range. Weighting the county-specific predictions for e by the number of households per county suggests weighted averages of .25 in SSBW and .33 in Tra. This comparison suggests consistency in structural and quasi-experimental measures of the average of the estimates for the marginal response to the policy. However, their implications for the benefits of the policy are quite different. The capitalization statistics measured by CG do not make it possible to evaluate how benefits differ across the metropolitan landscape as a result of changes in housing prices, changes in the hedonic price function, and changes in location choices. SSBW and Tra find that these effects matter for evaluating the distribution of benefits.

5.4.2 *Local Government Policy*

The ability to account for how heterogeneous households choose among a diverse choice set and how their subsequent actions contribute to equilibria is one of the strengths of the sorting framework. Indeed, the diverse features of households and how they contribute to the differences in outcomes across communities are what many policy makers are interested in learning about. Economists, especially those working on benefit cost analysis, often overlook the importance of these "details". It is not simply who will gain and lose, but will integration of neighborhoods increase or decrease? Will the efforts to improve air quality complement or undermine efforts to improve local public education? The former effect could be due to higher income households' willingness to support public education and the later due to housing price

⁵³ SSBW (2004) use a pure characteristics model based on Epple and Sieg (1999) whereas Tra (2010) uses a random utility model based on Bayer, McMillan, and Reuben (2004). SSBW include Los Angeles, Orange, Riverside, San Bernardino and Ventura counties. Tra (2010) omitted Ventura. SSBW also considered different community definitions taking advantage of the breakup of the LA school district. Tra's model also includes more controls for other local attributes than SSBW.

increases pushing out the lower income households where policy seeks to enhance educational outcomes.

When we consider other areas of policy analysis, aside from environmental rules, direct benefit measures are typically a first order concern. Ferreyra's summary of the outcomes of two voucher policies identifies over forty outcome measures—approximately one-fifth relate to the distribution of economic measures of the gains. The rest describe effects such as the changes in educational quality, which households go to particular schools, expenditures for education, and tax rates. These details might be incidental to a benefit cost analysis but they are not incidental to a policy maker's assessment.

5.4.3 *Shocks to the System*

Shortly before we submitted this manuscript, the social science directorate of the National Science Foundation invited proposals to design research that would investigate the economic effects of disasters. One need only perform a Google search on “Hurricane Katrina economics” to conclude that economists are good at evaluating shocks after the fact (for example see Jacob Vigdor 2008). How about an *ex ante* assessment? Sorting models may help to bound the equilibrium outcomes for some shocks. Suppose a sorting model includes a spatial amenity such as a forest. How can one measure the effects of proximity to the forest and its size on land values in the surrounding neighborhoods? Burn it down! That is, simulate the effect of removing it. This does not capture the disamenity created by the scorched earth—but it does offer a gauge of the new equilibrium and with calibration of the model and ways to address the possibility of multiple equilibria we might be able to do more.

6. *Research Frontiers*

Recent research has begun to extend sorting models in several ways. We conclude our review by considering five that are especially promising.⁵⁴

6.1 *Dynamics and Forward Looking Agents*

⁵⁴ Most of the papers we consider in this section are, at the time of writing, unpublished working papers.

Epple, Romano and Sieg (2010) extend the vertical class of sorting models to include overlapping generations in a multi-jurisdictional model of equilibrium voting. They introduce two important forms of heterogeneity that inject dynamics into the household's residential location decision—moving costs and age/family structure. Their model assumes that over the household's life-cycle, its preferences for education and housing services will evolve. Forward looking agents will recognize moving costs when making location choices early in life (i.e. child rearing years) as well as in their retirement years. Families with low moving costs will choose to move to areas with low education expenditures once their children depart, while those with strong preferences for public goods may choose to stay. These dynamics also affect the voting process, and with it the equilibrium levels of public goods. In other words, their model makes education expenditures and tax rates endogenous while accounting for evolving preferences over time.

Focusing on the class of stationary equilibria, Epple, Romano, and Sieg calibrate their model to the Boston metropolitan area. They find parameters that make the model predict expenditure shares from the Consumer Expenditure Survey (e.g. housing) and the fraction of people who move to a new location when they enter the “old” life stage. Mobility costs are found to play an important role in the evolution of communities. In a world without moving costs, more of the old would move to low education communities. Interestingly, by becoming older those communities would also become wealthier, raising their tax base and reducing the disparity between low and high education communities.

As we noted earlier, the sorting literature has either treated all households as renters (e.g. Sieg et al. 2004) or as myopic homeowners (e.g. Nechyba 1999, 2000). These strategies ignore an important dimension of the home-buying decision. Households will purchase a house in a neighborhood where they expect to receive a return on their investment that is consistent with comparable alternatives. Until recently, these returns were regarded as relatively secure. Increasing property values provided capital gains for homeowners. While the literature has yet to model the choice between owning and renting, Bayer et al. (2011) develop an estimator that allows households to be forward looking with respect to the appreciation of their house values. Moreover, their model accounts for moving costs. Because moving costs prevent repeated re-optimization, they cause homebuyers to be forward looking with respect to the evolution of

neighborhood characteristics (e.g., crime, pollution, and race). This has important consequences for recovering household preferences.

Consider a simple example with two houses—the first is located in a high pollution, high crime, but improving neighborhood. The second is located in a low pollution, low crime, but deteriorating neighborhood. The forward looking household may be willing to pay a premium to live in the first neighborhood. In a static model where its decision is defined in terms of current neighborhood attributes, an estimate for this household’s marginal value of reducing crime or pollution will be biased downward. The opposite logic applies to the household choosing to live in the declining neighborhood.

The challenge in modeling dynamics arises from the size of the state space. Even in the conservatively parameterized model used in Bayer et al. (2011), individuals choose between houses in 225 neighborhoods described by five attributes (housing price, violent crime, ground-level ozone, and percentage white). Suppose each of these attributes is allowed to take on one of ten values. The result is 10^{1125} potential points in the state space that might be visited in the future. This is impractical using traditional computational methods for estimating dynamic decision processes (John Rust 1987). Models of dynamic demand for consumer durables in industrial organization have dealt with a similar state-space problem by assuming that the logit inclusive value of the choice set is sufficient to describe its value, and that this inclusive value evolves according to some statistical process, such as first-order autoregressive scheme (Oleg Melnikov 2001; Juan Esteban Carranza 2007; Gautam Gowrisankaran and Marc Rysman 2007; Pasquale Schiraldi 2011). This reduces the size of the state space to one.

Bayer et al. (2011) avoid this assumption by instead employing a variation of V. Joseph Hotz and Robert A. Miller’s (1993) two-step approach to dynamic optimization. Estimates of choice-specific value functions (net of moving costs) are first recovered directly from the observed decisions of movers. These value function estimates are then used in a second-stage to recover estimates of moving costs from the move-stay decision. By measuring the financial component of moving costs, with 6% of the sale price paid to a realtor, this stage of the estimation procedure also yields an estimate of the marginal utility of wealth. Combining the estimates of choice-specific value functions and moving costs within a Bellman representation, estimates of flow utilities associated with each neighborhood and wealth group are recovered in a third

stage. These are decomposed to recover utility parameters in a final stage. Comparing the results with a static sorting model suggests that failure to recognize forward-looking behavior biases estimates of marginal willingness to pay for reducing violent crime and ground-level ozone by 34% and 20% respectively. Conversely, the static estimate of whites' marginal willingness-to-pay for same racial composition (which exhibits tremendous persistence over time) is biased upward by 71%.

Whereas calibration in Epple, Romano, and Sieg (2012) is based on the calculation of a dynamic sorting equilibrium, Bayer et al. (2011) does not yield equilibrium predictions. Their strategy avoids the challenges posed by a large choice set and technical feasibility of solving the dynamic programming problem. It cannot therefore solve for a new price vector when there are exogenous changes and thus it is unable to replicate the type of policy analysis performed by Epple, Romano, and Sieg.

An important task for future research is to develop computationally tractable methods to calculate dynamic sorting equilibria while simultaneously preserving the richness of dynamic decision making. One possibility, described in Peter Arcidiacono et al. (2010) and related work by Ulrich Dorazelski and Kenneth Judd (2010), is a continuous time model. Viewing the sorting equilibrium as a high-dimensional multi-agent game, that game becomes hard to solve due to so many players making simultaneous moves. As the number of players grows, the size of the state space becomes prohibitive.⁵⁵ Modeling the problem in continuous time allows the researcher to treat decisions as if they occur sequentially. In other words, a random process determines if an agent is eligible to make a decision at each point in time, but no two players are allowed to move at the same time. The number of alternatives underlying the individual's expectation then grows multiplicatively. Continued progress along these lines may make it feasible to calculate counterfactuals for dynamic sorting equilibria. A dynamic model with forward looking behavior and moving costs could be used to estimate structural parameters describing the decision of whether to rent or own.

6.2. Agent Based Models

⁵⁵ In particular, in order to determine their optimal policy, players have to form an expectation over all the potential actions that every other agent might take. The number of alternatives underlying this expectation grows exponentially in the number of agents and choices.

Agent based models (ABM's) have largely been developed outside of the economics discipline, but are often used in the geographic sciences to answer questions similar to those posed in the sorting literature. Irwin (2010) summarizes this literature, highlighting the relationship to structural economic models. ABM's take a "bottom-up" approach, modeling interactions between heterogeneous agents by specifying a set of behavioral rules and protocols—in other words, an entire institutional infrastructure in which agents interact. Simulated interactions between agents yield aggregate outcomes that can be matched to macro data for purposes of model calibration. These models generally allow for detailed spatial heterogeneity and dynamics, where an outcome at one location can influence the evolution of state variables at another.

Whereas sorting models make predictions about equilibrium outcomes, the features of agent-based models make them better suited to investigating the transitional dynamics *between* equilibria. This may prove to be useful if re-equilibration is a slow process relative to the time-horizon of the policy-maker. A lack of equilibrium restrictions allows ABM's to more easily incorporate rich spatial and agent heterogeneity along with multiple feedback effects. These can be particularly important when studying the impacts of non-marginal policy changes. For example, an ABM could be used to characterize the "basin of attraction" for combinations of parameter values and policy variables that lead to desirable or undesirable tipping effects.

The current generation of agent based models is limited because they rarely invoke the assumptions of optimizing behavior that underlie the equilibrium sorting literature. Agents need not be rational from either a static or dynamic perspective. Moreover, ABM's do not typically require the collective actions of individuals to satisfy basic market clearing conditions. Thus, an ABM simulation may produce a much wider set of potential outcomes than a structural model that imposes the discipline required by economic constraints. *Some* flexibility seems desirable, but allowing all behavior to be random with simple survival rules determining overall outcomes and a limited role for prices seems to ignore what we have learned about the functioning of markets. A blend of systematic rational responses at the agent level and less discipline imposed by static (or dynamic) market equilibrium conditions would seem a promising path for future research.

6.3. *Housing Supply*

Recent work has sought to develop a more realistic portrayal of supply-side behavior in housing markets. Consider, for example, models of suburban sprawl and open space preservation. Epple, Gordon, and Sieg (2010), citing Glaeser, Joseph Gyorko, and Raven Saks (2005), note that the elasticity of housing supply is of crucial importance in determining how cities respond to macroeconomic shocks. Are these shocks transmitted to prices and wages, or do they simply result in more sprawl?

It is challenging to estimate a housing supply function. We observe transactions that reflect newly constructed houses and sales of pre-existing structures. The former reflect past decisions of builders, due to a relatively long production cycle, and the later reflect more recent decisions of current owners. In this case there is a selection effect. We observe what current owners choose to offer given their expectations for prices, not what would be offered under different expectations. Builders' decisions are analogous, but their expectations relate to an earlier period when construction must have begun in order to sell the house in the current period.⁵⁶ Given a constant returns to scale production function for housing, and variation in land prices, Epple, Gordon and Sieg (2010) demonstrate that the production function for housing and the supply of housing services can be recovered while treating both prices and quantities as latent variables. Nonparametric estimates are derived by specifying supply per unit of land.⁵⁷

While an important first step, there are several reasons to consider extensions. First Epple et al. (1984,1993) noted that the conditions for an equilibrium in multi-community models rule out arbitrarily small communities. If the choice alternatives in a sorting model were subdivisions (with homeowner associations voting on services) then we would want to reconsider the constant return to scale assumption. Equally important, Klaiber and Phaneuf (2010a) suggest that builder size, measured by total houses built, does matter. Their analysis is the first static, random utility version of supply designed for a sorting model. They highlight three challenges: (i) identifying builders and their attributes using housing transactions alone; (ii) understanding and modeling differences in the choice set for builders versus those demanding home/community

⁵⁶ Of course, houses are also built for buyers. We are ignoring this issue here and considering houses built without specific buyers identified at the time of the construction decision.

⁵⁷ This strategy is reminiscent of an early argument in the hedonic literature by Parsons (1990) to normalize housing prices per unit of land.

combinations and; (iii) recovering estimates of the alternative specific constants with limited sample coverage of elements in the choice set. This can cause the estimates to be sensitive to outliers.

Murphy (2010) extends the static approach to treat builders as forward-looking agents, with a focus on their decisions about *when* to build. He notes that, while housing prices and quantities exhibit a high degree of volatility, they tend to move in cycles that exhibit strong serial correlation. In Murphy's model, a developer who owns a single plot of land decides each period whether to develop or wait, given his expectations about the evolution of future prices. If he chooses to develop, he decides on a level of quality. Using data from the California Bay Area in the 1990's, Murphy finds that many builders chose to build as prices begin a long (and predictable) run-up. There are two explanations for this seemingly odd behavior: (1) builders are static optimizers who react to the first sign of a price increase (an assumption that would be hard to justify on institutional grounds) or (2) there are pro-cyclical components of costs (e.g. fixed costs of permitting) and other building costs (e.g. difficulties in securing contractors) to which forward looking builders respond in their timing decisions. Murphy constructs a model consistent with the second explanation and uses it to recover estimates of these costs. He finds that pro-cyclical building costs can explain why it is in builders' interests to smooth out construction patterns over time. Moreover, these costs are correlated with high rates of home ownership in the cross-section, suggesting a political economy story in which existing residents play a role in limiting supply (John M. Quigley 2006; Kahn 2007; François Ortalo-Magne and Andrea Prat 2005).

Research on the political economy of housing supply has focused on the potential efficiency gains from zoning. The potential for inefficiency arises because poor households have an incentive to build small houses in rich communities where their property tax burden is less than their contribution to the cost of providing the public good. Faced with this problem, Bruce W. Hamilton (1975) reasoned that a law defining the minimum house size could ensure the efficiency of Tiebout sorting. Michelle J. White (1975) considered the ways in which homeowners could design zoning laws to ensure that incoming residents pay their marginal cost of public services, or be "squeezed" to pay even more. Fernandez and Rogerson (1997) and Calabrese, Epple, and Romano (1997) developed static models of multi-community equilibria with voting on both zoning laws and a property tax used to fund a public good, and Stephen Coate (2011)

recently developed a dynamic model. Theoretical results and numerical examples from these three studies suggest that zoning has the *potential* to improve efficiency but the gains may be unevenly distributed and local governments may fail to choose the most efficient regulations.

There is considerable scope for further supply side research. For example, the literature has yet to consider how large developers (who are capable of exercising market power) might alter supply relationships. One could also extend the current political economy models to consider the implications of zoning restrictions for the supply of endogenous amenities that are not funded from property taxes, such as open space, commuting patterns, and peer effects in public schools. It would also be interesting to endogenize zoning in an estimable model of sorting behavior. Calabrese et al. (2006) provide a blueprint for how this might be done.

6.4. Labor Supply

For working households, there are two dimensions of location choice—the choice of a house and the choice of a job. Representing both choices as part of a “dual-market” sorting model could improve our understanding of preferences for amenities and the evolution of neighborhoods. This suggestion is underscored by Paul W. Rhode and Koleman S. Strumpf’s (2003) historical assessment of Tiebout sorting. As moving costs declined between 1850 and 1990, they find that U.S. counties and municipalities became *less* stratified by public goods provision and household demographics; the opposite of what we would predict in a traditional static model of Tiebout sorting (i.e. holding job locations fixed). Furthermore, 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.

Rosen (1979) outlined the conceptual logic for dual-market sorting. Because households are free to adjust their behavior in both markets, we should expect both wage rates and house prices to reflect spatial variation in amenities. Intuitively, locations providing fewer amenities must offer higher wages and lower housing prices to induce households to locate there. As more households locate in high-amenity areas, the supply of labor and the demand for housing will increase, decreasing wages and increasing rents. Thus, households “pay” for the amenities in their area through both rents and wages. Jennifer Roback (1982) formalized Rosen’s intuition. In her model of interregional sorting, a household’s implicit expenditures on amenities are jointly

determined by a hedonic wage function and a hedonic price function. Empirical applications have reported that wage differentials reflect a substantial share of the total expenditures on amenities (Glenn C. Blomquist, Mark C. Berger, and John P. Hoehn 1988, Philip E. Graves and Donald M. Waldman 1991, Kahn 1995; and David Albouy 2009).⁵⁸

Most of the literature following Rosen (1979) and Roback (1982) has yet to incorporate the insights from the sorting literature. Households are typically assumed to *freely sort across the nation* based on their *homogeneous preferences for exogenous amenities*. Each metropolitan area offers a different (price, wage, amenity) bundle, but there is assumed to be *no heterogeneity in the spatial landscape* within a metro area.

Bayer, Keohane, and Timmins (2009) relax the national free mobility assumption. They find that indirectly controlling for the cost of moving between metro areas triples their estimates of the willingness-to-pay for a small improvement in air quality. While their empirical model tracks changes that occur over time, they do not model forward looking behavior. Other analysts have sought to adapt the recent advances in modeling dynamics. John Kennan and James R. Walker (2010) model interstate migration decisions with the goal of describing the role of expected future income prospects. This exercise is complicated by the need to account for the possibility of repeat and return migration decisions. In other words, they model optimal *sequences* of migration decisions, instead of a one-time decision to move or not move. Relying on numerical discrete approximation techniques to calculate the value function (Rust 1994), Kennan and Walker define the choice set as U.S. states, described only by time-invariant population and climate. Moreover, to shrink the state space, they restrict individuals' information sets to include only wages seen in recently visited locations. In a similar model of family location decisions accounting for marital status, Ahu Gemici (2008) restricts the choice set to nine US census regions. While both papers take an important step toward modeling forward looking behavior, their choice sets and data are not practical for analyzing heterogeneous policy outcomes at a high level of spatial resolution.

Bishop (2009) adapts the two-step approach described by Arcidiacono and Miller (2011) to the Kennan-Walker interstate sorting context and proposes a forward looking model for

⁵⁸ Roback's model is most commonly used to calculate "quality-of-life" indices. These indices rank metro areas by the implicit cost of consuming their bundle of amenities.

analyzing outcomes that vary across metropolitan areas. Arcidiacono and Miller's (2011) approach reduces Kennan and Walker's computational burden by treating the choice-specific value function as a sequence of flow utilities, conditional choice probabilities (CCP's), and the present discounted value of the value function realized at some future state.⁵⁹ A limited memory assumption can then be used to difference away that future value function.⁶⁰ This strategy converts a complex dynamic decision problem to a simple discrete choice over combinations of differenced flow utilities and CCP's. Utility parameters are then estimated with simple discrete choice techniques (e.g. multinomial logit), allowing any number of time-varying attributes without encountering the curse of dimensionality.

Finally, Kuminoff (2010) returns to a static setting to characterize dual-market sorting at an even higher level of spatial resolution. He focuses on the San Francisco and Sacramento metro areas because they contain diverse housing communities within several cities, which he treats as distinct job locations. Kuminoff's dual-market framework extends the structural model from Epple and Sieg (1999) in two ways. First, wage income and leisure time are both endogenous to location choice. Working households with heterogeneous job skills are assumed to select a job-house combination based on the wages they can earn, their preferences for the amenities provided by housing communities, 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 amenities. Thus, in a dual-market locational equilibrium, working households are simultaneously sorted among housing and labor markets according to their heterogeneous preferences and skills. Opportunities for adjustment in both markets make the implicit cost of consuming amenities depend on housing prices, wage rates, and commute times.

Spatial variation in labor and environmental regulations has recently been shown to affect the location choices of manufacturing firms as well as the supply and demand for labor (e.g. Thomas J. Holmes 1998, Kahn and Erin Mansur 2010, W. Reed Walker 2012). A dual-market model of sorting equilibrium could provide a basis for assessing the distributional implications of

⁵⁹ Hotz and Miller (1993) originally described the two-step method for solving complex dynamic programming problems by using conditional choice probabilities to approximate value functions.

⁶⁰ This assumption exploits the idea that an individual will have the same present discounted value of future utility starting from some point in the future regardless of the path taken to get there if her memory (i.e., the extent to which current utility is dependent upon previous decisions, aside from their effect on the current state) is limited. For example, if an individual only remembers her wage draw from the last location visited in a sequence of migration decisions, the value function associated with being in a particular state in period $t+2$ will not be a function of the location choice taken in period t .

prospective changes in regulatory structure that would alter the quality of amenities, the demand for labor, and the location choices made by working households.

6.5. Model Validation

Over the past decade, there has been considerable progress toward relaxing the least credible assumptions maintained in early sorting models, such as homogenous preferences and exogenous amenities. Nonetheless, the current structural estimators still rely on parametric assumptions for utility functions (CES, quasi-linear), assumptions for the statistical distributions used to characterize sources of unobserved heterogeneity (log-normal, Type I extreme value), and assumptions eliminating sources of market friction. Many of the current estimators also rely on identification strategies that have not exploited the new “best practices” for mitigating omitted variable bias. Future studies should address these limitations by continuing to build on the insights from the recent quasi-experimental literature on how to locate exogenous variation in amenities.⁶¹

One approach would be to refine the current estimators using insights from the econometric literature on partial identification (see Manski 2007 and Elie Tamer 2010 for summaries). The general idea is to characterize the potential sensitivity of outcome measures to the least credible assumptions needed to obtain point estimates for model parameters. The resulting bounds on welfare measures, capitalization effects, or migration patterns may or may not be informative for policymakers. Likewise, the observable variation in a credible instrument may or may not be capable of providing informative bounds on a policy-relevant parameter. It will depend on the data and the question at hand. The point is that moving from point identification to partial identification would allow us to provide a more transparent characterization of the mapping between assumptions and inferences.⁶² Banzhaf and Smith (2007) and Kuminoff (2009) provide two illustrations of how this approach can be used to investigate the extent to which estimates for structural preference parameters are identified by variations in the data relative to maintained assumptions about household preference functions and choice sets.

⁶¹ Bayer, Ferreira, and McMillan (2007) and Bayer, Keohane, and Timmins (2009) provide examples.

⁶² While many studies have conducted extensive sensitivity analyses (e.g. Bayer, Ferreira and McMillan 2007, Sieg et al. 2004), partial identification would provide a more formal and more comprehensive framework for evaluating the sensitivity of inferences to subjective modeling decisions.

The next step is to look for ways to narrow the set of candidate models. Keane (2010) emphasizes the need to perform validation exercises as part of the process. One possibility is to use the models to predict observable outcomes that are not used as fitting criteria during the estimation. Galiani, Murphy, and Pantano (2012) provide the first application of this approach. They estimate an RU model of sorting in the Boston metropolitan area and use it to make out-of-sample predictions for the distribution of households that would choose to participate in a new subsidy program designed to help low-income households in low-income neighborhoods move to higher-income neighborhoods. The model is found to be quite accurate in its predictions for participation in the “Moving-to-Opportunity” program operated by the U.S. Department of Housing and Urban Development. Another possibility would be to compare a sorting model’s prediction for how markets would respond to a shock with quasi-experimental evidence on how the same market did respond to an actual shock. We provided an example of this approach in section 5.4.1 when we compared Chay and Greenstone’s (2005) evidence on the elasticity of housing expenditures to air quality with the predictions implied by Sieg et al. (2004) and Tra (2010). These and other approaches to model validation will be needed to judge future advances in equilibrium models of sorting behavior.

REFERENCES

- Albouy, David. 2009. “What are Cities Worth? Land Rents, Local Productivity and the Capitalization of Amenity Values.” *NBER Working Paper* #14981.
- Alonso, William. 1964. *Location and Land Use: Towards a General Theory of Land Rent*. Cambridge: Cambridge University Press.
- Anderson, Simon P., André de Palma, and Jaques-François Thisse. 1992. *Discrete Choice Theory of Product Differentiation*. Cambridge and London: The MIT Press.
- Angrist, Joshua D. and Jörn-Steffen Pischke. 2010. “The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics.” *Journal of Economic Perspectives*, 24(2): 3-30.
- Arcidiacono, Peter and Robert Miller. 2011. “CCP Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity.” *Econometrica*, Vol. 79(6): 1823-1867.
- Arcidiacono, Peter, Patrick Bayer, Jason Blevins, and Paul Ellickson. 2010. “Estimation of Dynamic Discrete Choice Models in Continuous Time.” *ERID Working Paper No.50*.

- Bajari, Patrick and C. Lanier Benkard. 2005. "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach." *Journal of Political Economy*, 113(6): 1239-76.
- Bajari, Patrick, Han Hong, John Krainer, and Denis Nekipelov. 2010. "Estimating Static Models of Strategic Interaction." *Journal of Business and Economic Statistics*, Vol. 28(4): 469-482.
- Bajari, Patrick and Matthew E. Kahn. 2005. "Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities." *Journal of Business and Economic Statistics*, 23(1): 20-33.
- Banzhaf, H. Spencer and V. Kerry Smith. 2007. "Meta-analysis in Model Implementation: Choice Sets and the Valuation of Air Quality Improvements." *Journal of Applied Econometrics*, 22(6): 1013-31.
- Banzhaf, H. Spencer and Randall P. Walsh. 2008. "Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism." *American Economic Review*, 98(3): 843-63.
- Banzhaf, H. Spencer and Randall P. Walsh. 2010. "Segregation and Tiebout Sorting: Investigating the Link between Investments in Public Goods and Neighborhood Tipping." *NBER Working Paper #16057*.
- Bartik, Timothy J. 1987. "The Estimation of Demand Parameters in Hedonic Price Models." *Journal of Political Economy*, 95(1): 81-88.
- Bartik, Timothy J. 1988. "Measuring the Benefits of Amenity Improvements in Hedonic Price Models." *Land Economics*, 64(2): 172-183.
- Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy*, 115(4): 588-638.
- Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. "Migration and Hedonic Valuation: The Case of Air Quality." *Journal of Environmental Economics and Management*, 58(1).
- Bayer, Patrick, Robert McMillan, Alvin Murphy, and Christopher Timmins. 2011. "A Dynamic Model of Demand for Houses and Neighborhoods." *NBER Working Paper #17250*.
- Bayer, Patrick, Robert McMillan, and Kim Reuben. 2004. "An Equilibrium Model of Sorting in an Urban Housing Market." *NBER Working Paper No 10865*.
- Bayer, Patrick, Steve Ross, and Giorgio Topa. 2008. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes." *Journal of Political Economy*, 116(6): 1150-1196.
- Bayer, Patrick and Christopher Timmins. 2005. "On the Equilibrium Properties of Locational Sorting Models." *Journal of Urban Economics*, 57(3): 462-77.
- Bayer, Patrick and Christopher Timmins. 2007. "Estimating Equilibrium Models of Sorting across

- Locations." *The Economic Journal*, 117(518): 353-74.
- Benabou, Roland. 1993. "Workings of a City: Location, Education, and Production." *Quarterly Journal of Economics*, 108(3): 619-52.
- Benabou, Roland. 1996a. "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance." *American Economic Review*, 86(3): 584-609.
- Benabou, Roland. 1996b. "Equity and Efficiency in Human Capital Investment: The Local Connection." *Review of Economic Studies*, 63(2): 237-64.
- Berry, Steven and Ariel Pakes. 2007. "The Pure Characteristics Demand Model." *International Economic Review*, 48(4): 1193-225.
- Berry, Steven, Oliver B. Linton, and Ariel Pakes. 2004. "Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems." *Review of Economic Studies*, 71(3): 613-54.
- Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841-90.
- Bergstrom, Theodore C. and Robert P. Goodman. 1973. "Private Demands for Public Goods." *American Economic Review*, 63(3): 280-96.
- Bergstrom, Theodore C., Daniel L. Rubinfeld, and Perry Shapiro. 1982. "Micro-Based Estimates of Demand Functions for Local School Expenditures." *Econometrica*, 50(5): 1183-205.
- Bishop, Kelly. 2009. "A Dynamic Model of Location Choice and Hedonic Valuation." *Working Paper*.
- Black, Sandra E. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics*, 114(2): 577-99.
- Blinder, Alan S., and Harvey S. Rosen. 1985. "Notches." *American Economic Review*, 75(4): 736-47.
- Blomquist, Glenn C., Mark C. Berger, and John P. Hoehn. 1988. "New Estimates of Quality of Life in Urban Areas." *American Economic Review*, 78(1): 89-107.
- Bresnahan, Timothy F. 1987. "Competition and Collusion in the American Automobile Industry: The 1955 Price War." *Journal of Industrial Economics*, 35(4): 457-82.
- Brown, James N. and Harvey S. Rosen. 1982. "On the Estimation of Structural Hedonic Price Models." *Econometrica*, 50(3): 765-68.
- Bullock, David S. and Nicholas Minot. 2006. "On Measuring the Value of a Nonmarket Good Using Market Data." *American Journal of Agricultural Economics*, 88(4): 961-73.
- Calabrese, Stephen, Dennis Epple, and Richard Romano. 2007. "On the Political Economy of Zoning." *Journal of Public Economics*, 97(1-2): 25-49.
- Calabrese, Stephen, Dennis Epple, Thomas Romer, and Holger Sieg. 2006. "Local Public Good Provi-

- sion: Voting, Peer Effects, and Mobility.” *Journal of Public Economics*, 90(6-7): 959-981.
- Card, David, Alexandre Mas, and Jesse Rothstein. 2008. “Tipping and the Dynamics of Segregation.” *Quarterly Journal of Economics*, 123(1): 177-218.
- Carranza, Juan Esteban. 2007. “Estimation of Demand for Differentiated Durable Goods.” *Working Paper*.
- Chattopadhyay, Sudip. 1999. "Estimating the Demand for Air Quality: New Evidence Based on the Chicago Housing Market." *Land Economics*, 75(1): 22-38.
- Chay, Kenneth Y. and Michael Greenstone. 2005. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy*, 113(2): 376-424.
- Coate, Stephen. 2011. “Property Taxation, Zoning, and Efficiency: A Dynamic Analysis.” *NBER Working Paper #17145*.
- Cornes, Richard. and Roger Hartley. 2007. “Aggregative Public Good Games.” *Journal of Public Economic Theory*, 9(2): 201-219.
- Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor. 1999. "The Rise and Decline of the American Ghetto." *Journal of Political Economy*, 107(3): 455-506.
- Davis, Lucas. 2004. "The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster." *American Economic Review*, 94(5): 1693-704.
- De Bartolome, Charles A.M. 1990. "Equilibrium and Inefficiency in a Community Model with Peer Group Effects." *Journal of Political Economy*, 98(1): 110-33.
- Diewert, Walter .E. 1993. “The Early History of Price Index Reserch.” in *Essays in Index Number Theory, Volume 1*. W. Erwin Diewert and Alice O. Nakamura eds. Amsterdam: North Holland, pp. 33-66.
- Doraszelski, Ulrich and Kenneth Judd. 2008. “Avoiding the Curse of Dimensionality in Dynamic Stochastic Games.” *Working Paper*.
- Driscoll, Paul, Brian Dietz, and Jeffrey Alwang. 1994. "Welfare Analysis When Budget Constraints Are Nonlinear: The Case of Flood Hazard Reduction." *Journal of Environmental Economics and Management*, 26(2): 181-99.
- Durlauf, Steven N. 1996. “A Theory of Persistent Income Inequality.” *Journal of Economic Growth*, 1(1): 75-93.
- Ekeland, Ivar, James J. Heckman, and Lars Nesheim. 2004. "Identification and Estimation of Hedonic Models." *Journal of Political Economy*, 112(1): S60-S109.
- Ellickson, Bryan. 1971. "Jurisdictional Fragmentation and Residential Choice." *American Economic Review*, 61(2): 334-39.

- Epple, Dennis. 1987. "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products." *Journal of Political Economy*, 95(1): 59-80.
- Epple, Dennis and Maria Marta Ferreyra. 2008. "School Finance Reform: Assessing General Equilibrium Effects." *Journal of Public Economics*, 92(5-6): 1328-1351.
- Epple, Dennis, Brett Gordon, and Holger Sieg. 2010. "A New Approach to Estimating the Production Function for Housing." *American Economic Review*. 100(3):905-924.
- Epple, Dennis, Radu Filimon, and Thomas Romer. 1984. "Equilibrium Among Local Jurisdictions: Toward an Integrated Treatment of Voting and Residential Choice." *Journal of Public Economics*, 24(3): 281-308.
- Epple, Dennis, Radu Filimon, and Thomas Romer. 1993. "Existence of Voting and Housing Equilibria in a System of Communities with Property Taxes." *Regional Science and Urban Economics*, 23(5): 585-610.
- Epple, Dennis, Michael Peress, and Holger Sieg. 2010. "Identification and Semiparametric Estimation of Equilibrium Models of Local Jurisdictions." *American Economic Journal: Microeconomics*, 2(4): 195-220.
- Epple, Dennis and Glenn J. Platt. 1998. "Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Incomes." *Journal of Urban Economics*, 43(1): 23-51.
- Epple, Dennis, Richard Romano, and Holger Sieg. 2006. "Admission, Tuition, and Financial Aid Policies in the Market for Higher Education." *Econometrica*, 74(4): 885-928.
- Epple, Dennis, Richard Romano, and Holger Sieg. 2012. "The Intergenerational Conflict over the Provision of Public Education." *Journal of Public Economics*, in press.
- Epple, Dennis and Thomas Romer. 1991. "Mobility and Redistribution." *Journal of Political Economy*, 99(4): 828-58.
- Epple, Dennis and Holger Sieg. 1999. "Estimating Equilibrium Models of Local Jurisdiction." *Journal of Political Economy*, 107(4): 645-81.
- Ferreira, Fernando. 2010. "You Can Take it With You: Proposition 13 Tax Benefits, Residential Mobility, and Willingness to Pay for Housing Amenities." *Journal of Public Economics*, 94(9-10): 661-673.
- Fernandez, Raquel and Richard Rogerson. 1996. "Income Distribution, Communities, and the Quality of Public Education." *Quarterly Journal of Economics*, 111(1): 135-64.
- Fernandez, Raquel and Richard Rogerson. 1997. "Keeping People Out: Income Distribution, Zoning, and the Quality of Public Education." *International Economic Review*, 38(1): 23-42.

- Fernandez, Raquel and Richard Rogerson. 1998. "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform." *The American Economic Review*, 88(4): 813-833.
- Ferreira, Maria Marta. 2007. "Estimating the Effects of Private School Vouchers in Multi-District Economies." *American Economic Review*, 97(3): 789-817.
- Galiani, Sebastian, Alvin Murphy, and Juan Pantano, 2012, "Estimating Neighborhood Choice Models: Lessons from a Housing Assistance Experiment" unpublished paper, Washington University, June.
- Gamper-Rabindran, Shanti, and Christopher Timmins. 2011. "Does Cleanup of Hazardous Waste Sites Raise Housing Values? Evidence of Spatially Localized Benefits." *Working paper*.
- Gemici, Ahu. 2008. "Family Migration and Labor Market Outcomes." *Working Paper*, New York University.
- Glaeser, Edward L., Joseph Gyourko, and Raven Saks. 2005. "Urban Growth and Housing Supply." *Working Paper*.
- Goldstein, Gerald S. and Mark V. Pauly. 1981. "Tiebout Bias on the Demand for Local Public Goods." *Journal of Public Economics*, 16(2): 131-43.
- Gorman, William M. 1980. "A Possible Procedure for Analysing Quality Differentials in the Egg Market." *Review of Economic Studies*, 47(5): 843-856.
- Gowrisankaran Gautam and Marc Rysman. 2006. "Dynamics of Consumer Demand for New Durable Goods." *Working Paper*.
- Graves, Philip E. and Donald M. Waldman. 1991. "Multimarket Amenity Compensation and the Behavior of the Elderly." *American Economic Review*, 81(5): 1374-81.
- Greenstone, Michael and Justin Gallagher. 2008. "Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program." *Quarterly Journal of Economics*, 123(3): 951-1003.
- Griliches, Zvi. 1971. "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change," in *Price Indexes and Quality Change*. Zvi Griliches ed. Cambridge: Harvard University Press.
- Haefele, Edwin T. 1973. "A Utility Theory of Representative Government." *American Economic Review*, 61(3): 350-367.
- Hallstrom, Daniel G. and V. Kerry Smith. 2003. "Habitat Protection Policies and Open Space: A General Equilibrium Analysis of "Takings" and "Givings". *Working Paper*.
- Hamilton, Bruce W. 1975. "Zoning and Property Taxation in a System of Local Governments." *Urban Studies*, 12(2): 205-211.

- Hanushek, Eric A. 2003. "The Failure of Input-Based Schooling Policies." *The Economic Journal*, 113(485): F64-F98.
- Hansen, Lars Peter, and James J. Heckman. 1996. "The Empirical Foundations of Calibration." *Journal of Economic Perspectives*, 10(1): 87-104.
- Harberger, Arnold C. 1964. "The Measurement of Waste." *American Economic Review*, 54(3): 58-76.
- Hausman, Jerry A. 1997. "Valuation of New Goods Under Perfect and Imperfect Competition." in *The Economics of New Goods*. Timothy F. Bresnahan and Robert J. Gordon eds. Chicago: The University of Chicago Press, pp. 209-237.
- Heckman, James J. 2010. "Building Bridges between Structural and Program Evaluation Approaches to Evaluating Policy." *Journal of Economic Literature*, 48(2): 356-398.
- Heckman, James J., Rosa M. Matzkin, and Lars Nesheim. 2010. "Nonparametric Identification of Nonadditive Hedonic Models." *Econometrica*, 78(5): 1569-1592.
- Hendel, Igal and Aviv Nevo. 2006. "Measuring the Implications of Sales and Consumer Inventory Behavior." *Econometrica*, 74(6): 1637-73.
- Holmes, Thomas J. 1998. "The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders." *Journal of Political Economy*, 106(4): 667-705.
- Hotz, V. Joseph. and Robert A. Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *Review of Economic Studies*, 60:497-529.
- Imbens, Guido M. and Jeffrey M. Wooldridge, 2009, "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 47 (March): 5-86, also NBER working paper # 14251, August, 2008.
- Irwin, Elena G. 2010. "New Directions for Urban Economic Models of Land Use Change: Incorporating Spatial Dynamics and Heterogeneity." *Journal of Regional Science*. 50(1):65-91.
- Kahn, Matthew E. 1995. "A Revealed Preference Approach to Ranking City Quality of Life." *Journal of Urban Economics*, 38(2): 221-35.
- Kahn, Matthew E. 2007. "Environmentalism as a Determinant of Housing Supply Regulation." *Working Paper*.
- Kahn, Matthew E. and Erin T. Mansur. 2011. "Do Local Energy Prices and Regulation Affect the Geographic Concentration of Employment? A Border Pairs Approach." *Revision of NBER Working Paper#16538*.
- Kahn, Shulamit and Kevin Lang. 1988. "Efficient Estimation of Structural Hedonic Systems." *International Economic Review*, 28:1, pp. 157-66.
- Kanemoto, Yoshitsugu. 1988. "Hedonic Prices and the Benefits of Public Policies." *Econometrica*, 56(4):

981-89.

- Keane, Michael P. 2010. "Structural vs. Atheoretic Approaches to Econometrics." *Journal of Econometrics*, 156 (1): 3-20.
- Kennan, John and James R. Walker. 2010. "The Effect of Expected Income on Individual Migration Decisions." *Econometrica*, 79(1): 211-251.
- Klaiber, H. Allen and Daniel J. Phaneuf. 2010a. "Linking Developers and Homeowners in a Horizontal Sorting Framework: Implications for General Equilibrium Welfare Measurement." *Working Paper*.
- Klaiber, H. Allen and Daniel J. Phaneuf. 2010b. "Valuing Open Space in a Residential Sorting Model of the Twin Cities." *Journal of Environmental Economics and Management*, 60(2): 57-77.
- Klaiber, H. Allen, and V. Kerry Smith. 2012a "Quasi Experiments, Hedonic Models, and Estimating Tradeoffs for Local Amenities," *Land Economics* forthcoming (also NBER working paper #4957)
- Klaiber, H. Allen and V. Kerry Smith. 2012b. "General Equilibrium Benefit Analyses for Social Programs." *Journal of Benefit Cost Analysis*, March , 1-50.
- Kotchen, Matthew J. 2006. "Green Markets and Private Provision of Public Goods." *Journal of Political Economy*, 114(4): 816-834.
- Kuminoff, Nicolai V. 2009. "Decomposing the Structural Identification of Nonmarket Values." *Journal of Environmental Economics and Management*, 57(2): 123-39.
- Kuminoff, Nicolai V., and Abdul S. Jarrah. 2010. "A New Approach to Computing Hedonic Equilibria and Investigating the Properties of Locational Sorting Models." *Journal of Urban Economics* 67(3): 322-335.
- Kuminoff, Nicolai V. and Jaren C. Pope. 2011. "Rosen Bias in the Willingness to Pay for Nonmarket Goods." *Working Paper*.
- Kuminoff, Nicolai V. 2010. "Partial Identification of Preferences for Public Goods in a Dual-Market Sorting Equilibrium." *Working Paper*.
- Lancaster, Kelvin J. 1966. "A New Approach to Consumer Theory." *Journal of Political Economy*, 74(2): 132-57.
- Lancaster, Kelvin J. 1979. *Variety, Equity, and Efficiency*. New York: Columbia University Press.
- Lee, David S. and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(2): 281-355.
- Li, Shanjun. 2012. "Traffic Safety and Vehicle Choice: Quantifying the Arms Race on American Roads." *Journal of Applied Econometrics*, 27(1): 34-62.

- Linden, Leigh and Jonah E. Rockoff. 2008. "Estimates of the Impact of Crime Risk on Property Values from Megan's Laws." *American Economic Review*, 98(3): 1103-27.
- Manski, Charles F. 2007. *Identification for Prediction and Decision*. Harvard University Press, Cambridge.
- Matzkin, Rosa L. 2007. "Nonparametric Identification." In *Handbook of Econometrics, Volume 6B*, ed. James J. Heckman and Edward E. Leamer, 5307-68. Amsterdam and Oxford, North-Holland.
- McConnell, Virginia and Margaret Walls. 2005. "The Value of Open Space: Evidence from Studies of Non-Market Benefits." *Working Paper: Resources for the Future*.
- McFadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*. Paul Zarembka ed. New York: Academic Press.
- McFadden, Daniel. 1978. "Modelling the Choice of Residential Location," in *Spatial Interaction Theory and Planning Models*. A. Karlqvist, L. Lundqvist, F. Snickars and J. Weibull eds. Amsterdam: North Holland, pp. 75-96.
- Melnikov, Oleg. 2001. "Demand for Differentiated Durable Products: The Case of the US Computer Printer Market." *Working Paper*.
- Mendelsohn, Robert. 1985. "Identifying Structural Equations with Single Market Data." *Review of Economics and Statistics*, 67(3): 525-29.
- Mills, Edwin S. 1967. "An Aggregative Model of Resource Allocation in a Metropolitan Area." *American Economic Review*, 57(2):197-210
- Murphy, Alvin. 2010. "A Dynamic Model of Housing Supply." *Working Paper*.
- Muth, Richard F. 1969 *Cities and Housing*. University of Chicago Press, Chicago.
- Nechyba, Thomas, J. 1997. "Existence of Equilibrium and Stratification in Local and Hierarchical Tiebout Economies with Property Taxes and Voting." *Economic Theory*, 10(2): 277-304.
- Nechyba, Thomas, J. 1999. "School Finance Induced Migration and Stratification Patterns: The Impact of Private School Vouchers." *Journal of Public Economic Theory*, 1(1): 5-50.
- Nechyba, Thomas, J. 2000. "Mobility, Targeting, and Private-School Vouchers." *American Economic Review*, 90(1): 130-46.
- Nevo, Aviv. 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica*, 69(2): 307-42.
- Oates, Wallace E. 1969. "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis." *Journal of Political Economy*, 77(6): 957-71

- Ortalo-Magne, François. and Andrea Prat (2005). "The Political Economy of Housing Supply." *Working Paper*.
- Palmquist, Raymond B. 1984. "Estimating the Demand for the Characteristics of Housing." *Review of Economics and Statistics*, 66(3): 394-404.
- Palmquist, Raymond B. 1992. "Valuing Localized Externalities." *Journal of Urban Economics*, 31: 59-68.
- Palmquist, Raymond B. 2005a. "Weak Complementarity, Path Independence, and the Intuition of the Willig Condition." *Journal of Environmental Economics and Management*, 49(1): 103-15.
- Palmquist, Raymond B. 2005b. "Property Value Models," in *Handbook of Environmental Economics, Volume 2*. Karl-Göran Mäler and Jeffery Vincent eds. Amsterdam: North Holland Press.
- Parsons, George R. 1990. "Hedonic Prices and Public Goods: An Argument for Weighting Locational Attributes in Hedonic Regressions by Lot Size." *Journal of Urban Economics*, 27(3): 308-321.
- Pope, Jaren C. 2008. "Fear of Crime and Housing Prices: Household Reactions to Sex Offender Registries." *Journal of Urban Economics*, 64(3): 601-14.
- Quigley, John M. 2006. "Froth in the Silicon Valley Housing Market?" *The Economist's Voice*. 3(5), Article 1.
- Rhode, Paul W. and Koleman S. Strumpf. 2003. "Assessing the Importance of Tiebout Sorting: Local Heterogeneity from 1850 to 1990." *American Economic Review*, 93(5): 1648-77.
- Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy*, 90(6): 1257-78.
- Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*, 82(1): 34-55.
- Rosen, Sherwin. 1979. "Wage-Based Indexes of Urban Quality of Life," in *Current Issues in Urban Economics*. P. Mieszkowski and M. Straszheim eds. Baltimore: John Hopkins University Press.
- Rothbarth, E. 1941. "The Measurement of Changes in Real Income under Conditions of Rationing." *Review of Economic Studies*, 8(2): 100-107.
- Rothstein, Jesse M. 2006. "Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition among Jurisdictions." *American Economic Review*, 96 (4): 1333-1350.
- Roy, A.D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers (New Series)*, 3(2): 135-46.
- Rubinfeld, Daniel L., Perry Shapiro, and Judith Roberts. 1987. "Tiebout Bias and the Demand for Local Public Schooling." *Review of Economics and Statistics*, 69(3): 426-37.

- Rust, John. 1987. "Optimal Replacement of GME Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica*. 55(5):999-1033.
- Rust, John. 1994. "Structural Estimation of Markov Decision Processes." In Engle, R. and McFadden, D. (Editors), *Handbook of Econometrics*, Volume 4. Elsevier.
- Rysman, Marc. 2004. "Competition Between Networks: A Study of the Market for Yellow Pages." *Review of Economic Studies*, 71(2): 483-512.
- Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2(3): 180-212.
- Samuelson, Paul A. 1948. "Consumption Theory in Terms of Revealed Preferences." *Economica*, 15(60): 243-53.
- Schelling, Thomas C. 1969. "Models of Segregation." *American Economic Review*, 59(2): 488-93
- Schiraldi, Pasquale. 2011. "Automobile Replacement: A Dynamic Structural Approach." *Rand Journal of Economics*, 42(2): 266-291.
- Schmidheiny, Kurt. 2006. "Income Segregation and Local Progressive Taxation: Empirical Evidence from Switzerland." *Journal of Public Economics*, 90(3): 429-458.
- Scotchmer, Suzanne. 1985. "Hedonic Prices and Cost/Benefit Analysis." *Journal of Economic Theory*, 37(1): 55-75.
- Seim, Katja. 2006. "An Empirical Model of Firm Entry with Endogenous Product-type Choices." *Rand Journal of Economics*, 37(3): 619-40.
- Sethi, Rajiv and Rohini Somanathan. 2004. "Inequality and Segregation." *Journal of Political Economy*, 112(6): 1296-321.
- Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf, and Randy Walsh. 2002. "Interjurisdictional Housing Prices in Location Equilibrium." *Journal of Urban Economics*, 52(1): 131-53.
- Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf, and Randy Walsh. 2004. "Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods." *International Economic Review*, 45(4): 1047-77.
- Smith, V. Kerry and H. Spencer Banzhaf. 2004. "A Diagrammatic Exposition of Weak Complementarity and the Willig Condition." *American Journal of Agricultural Economics*, 86(2): 455-66.
- Smith, V. Kerry and Spencer Banzhaf. 2007. "Quality Adjusted Price Indexes and the Willig Condition." *Economics Letters*, 94(1): 43-48.
- Smith, V. Kerry, Holger Sieg, H. Spencer Banzhaf, and Randy Walsh. 2004. "General Equilibrium Benefits for Environmental Improvements: Projected Ozone Reductions under EPA's Prospective Analysis for the Los Angeles Air Basin." *Journal of Environmental Economics and Management*,

- 47(3): 559-84.
- Takeuchi, Akie, Maureen Cropper, and Antonio Bento. 2008. "Measuring the Welfare Effects of Slum Improvement Programs: The Case of Mumbai." *Journal of Urban Economics*. 64(1): 65-84.
- Tamer, Elie. 2010. "Partial Identification in Econometrics." Forthcoming in *Annual Review of Economics*.
- Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy*, 64(5): 416-24.
- Timmins, Christopher. 2007. "If You Cannot Take the Heat, Get Out of the Cerrado...Recovering the Equilibrium Amenity Cost of Nonmarginal Climate Change in Brazil." *Journal of Regional Science*, 47(1): 1-25.
- Timmins, Christopher and Jennifer Murdock. 2007. "A Revealed Preference Approach to the Measurement of Congestion in Travel Cost Models." *Journal of Environmental Economics and Management*, 53(2): 141-290.
- Tinbergen, Jan. 1959. "On the Theory of Income Distribution," in *Selected Papers of Jan Tinbergen*. L.H. Klaassen, L.M. Koych and H.J. Witteveen eds. Amsterdam: North Holland.
- Tra, Constant I., 2010. "A Discrete Choice Equilibrium Approach to Valuing Large Environmental Changes." *Journal of Public Economics*, 94 (1-2): 183-196.
- Train, Kenneth E. 2003. *Discrete Choice Methods with Simulation*. Cambridge: Cambridge University Press.
- Triplett, Jack E. 1969. "Automobiles and Hedonic Quality Measurement." *Journal of Political Economy*, 77(May-June): 408-417.
- Triplett, Jack E. 1983. "Concepts of Quality in Input and Output Price Measures: A Resolution of the User Value-Resource Cost Debate." in *The U.S. National Income and Product Accounts: Selected Topics*, Murray F. Foss ed, pp. 269-311. NBER Studies in Income and Wealth, Volume 47. Chicago: University of Chicago Press.
- US Environmental Protection Agency. 1999. *The Benefits and Costs of the Clean Air Act 1990 to 2010*, Report to Congress.
- US Environmental Protection Agency. 2010. *Guidelines for Preparing Economic Analyses*, EPA 240-R-10_001.
- Varian, Hal R. 1982. "The Nonparametric Approach to Demand Analysis." *Econometrica*, 50(4): 945-74.
- Vigdor, Jacob. 2008. "The Economic Aftermath of Hurricane Katrina." *Journal of Economic Perspectives*, 22(4): 135-54.
- Walker, W. Reed. 2012. "The Transitional Costs of Sectoral Reallocation: Evidence from the Clean Air Act and the Workforce." *Working Paper*.

- Walsh, Randall L. 2007. "Endogenous Open Space Amenities in a Locational Equilibrium." *Journal of Urban Economics*, 61(2): 319-44.
- Waugh, Frederick V. 1968. "Quality as a Determinant of Vegetable Prices." *Studies in History, Economics and Public Law*. Number 312, First AMS Edition (from the Columbia University 1929 edition), Faculty of Political Science Columbia University eds. New York: AMS Press.
- Westoff, Frank. 1977. "Existence of Equilibria in Economies with a Local Public Good." *Journal of Economic Theory*, 14(1): 84-112.
- White, Michelle J. 1975. "Fiscal Zoning in Fragmented Metropolitan Areas." in *Fiscal Zoning and Land Use Controls*, Edwin S. Mills and Wallace E. Oates ed, pp. 31-100. Lexington: Lexington Books.
- Wu, JunJie and Seong-Hoon Cho. 2003. "Estimating Households' Preferences for Environmental Amenities Using Equilibrium Models of Local Jurisdictions." *Scottish Journal of Political Economy*, 50(2): 198-206.
- Yinger, John. 1976. "Racial Prejudice and Racial Residential Segregation in an Urban Model." *Journal of Urban Economics*, 3(4): 383-396.

TABLE 1
FEATURES OF EMPIRICAL HEDONIC AND SORTING MODELS USED FOR POLICY EVALUATION

		(1)	(2)	(3)	(4)	(5)	(6)
		Multi-Market Hedonic	Structural Hedonic	Nonparametric Hedonic	Pure Characteristics Sorting	Random Utility Sorting	Nechyba-Ferreya Sorting
Selected references	econometric methodology	Rosen (1974) Brown-Rosen (1982) Epple; Bartik (1987) Kahn-Lang (1988)	Driscoll et al. (1994)	Ekeland et al (2004) Heckman et al (2010)	Epple-Sieg (1999) Sieg et al. (2002, 2004)	Bayer et al. (2004) Bayer-Timmins (2005,2007)	Ferreya (2007)
	calibrated models		Klaiber-Smith (2009) Kuminoff-Jarrah (2010)		Epple-Platt (1998)		Nechyba (2000)
	estimated models	Palmquist (1984)	Chattopadhyay (1999) Bajari-Kahn (2005)		Smith et al. (2004) Walsh (2007)	Bayer et al. (2007) Timmins (2007) Klaiber-Phaneuf (2010)	Ferreya (2007)
Definition for the choice set	choice	house	house	House	house	house job	house, school, vote
	budget constraint	yes	yes	Yes	yes	no*	yes
	amenities may vary discretely	no	no	No	yes	yes	yes

TABLE 1 (CONTINUED)
 FEATURES OF EMPIRICAL HEDONIC AND SORTING MODELS USED FOR POLICY EVALUATION

		(1) Multi-Market Hedonic	(2) Structural Hedonic	(3) Nonparametric Hedonic	(4) Pure Characteristics Sorting	(5) Random Utility Sorting	(6) Nechyba-Ferreya Sorting
Restrictions on preferences	must specify utility function	no	yes	no, given weak separability of marginal utility	yes	yes	yes
	tastes for amenities	fully explained by demographics	horizontally differentiated	single amenity or vertical structure	vertically differentiated	horizontally differentiated	horizontally differentiated
	key distributional assumption	none	none	normalization of distribution for idiosyncratic error	lognormal bivariate distribution for income, preferences	iid type I EV shocks for every location	iid type I EV shocks for every location
Strategy for addressing endogenous amenities	instruments arising from model structure	none	none	none	income rank	substitute attributes	production function
	instruments used in applications to date	none	none	no known applications	income rank	substitute attributes + amenity discontinuity	production function
Issues for evaluating implications of large scale shocks	can solve for new market equilibrium	no	yes	no	yes	yes	yes
	data describe entire market	no	no	no	yes	no	yes
	# substitutes per location				2	<i>J-1</i>	<i>J-1</i>

* It is possible to relax this assumption. See section 4.4.1 for details.

TABLE 2
SORTING ESTIMATES FOR MARGINAL AND TOTAL WILLINGNESS FOR AIR QUALITY IMPROVEMENTS

Geographic Aggregate (County)	Sieg et al. (2004)					Tra (2010)				
	Δq	<i>MWTP</i>	<i>e</i>	<i>WTP_{PE}</i>	<i>WTP_{GE}</i>	Δq	<i>MWTP</i>	<i>e</i>	<i>WTP_{PE}</i>	<i>WTP_{GE}</i>
Los Angeles	.208	50	.277	1472	1556	.21	50	.321	589	746
Orange	.180	44	.268	901	1391	.18	42	.278	595	748
Riverside	.207	25	.114	834	372	.16	51	.451	628	805
San Bernardino	.163	24	.135	738	367	.16	57	.400	489	633
Ventura	.062	26	.419	164	725	-	-	-	-	-

Note: The two studies describe the annualized marginal willingness to pay (*MWTP*), partial equilibrium willingness to pay (*WTP_{PE}*) and general equilibrium willingness to pay (*WTP_{GE}*) for reductions in ozone concentrations (Δq) in the Los Angeles metropolitan area. Sieg et al. use a pure characteristics sorting model. Tra uses a random utility sorting model. We used their predictions for *WTP_{GE}* to calculate the elasticity of housing expenditures to changes in air quality (*e*).

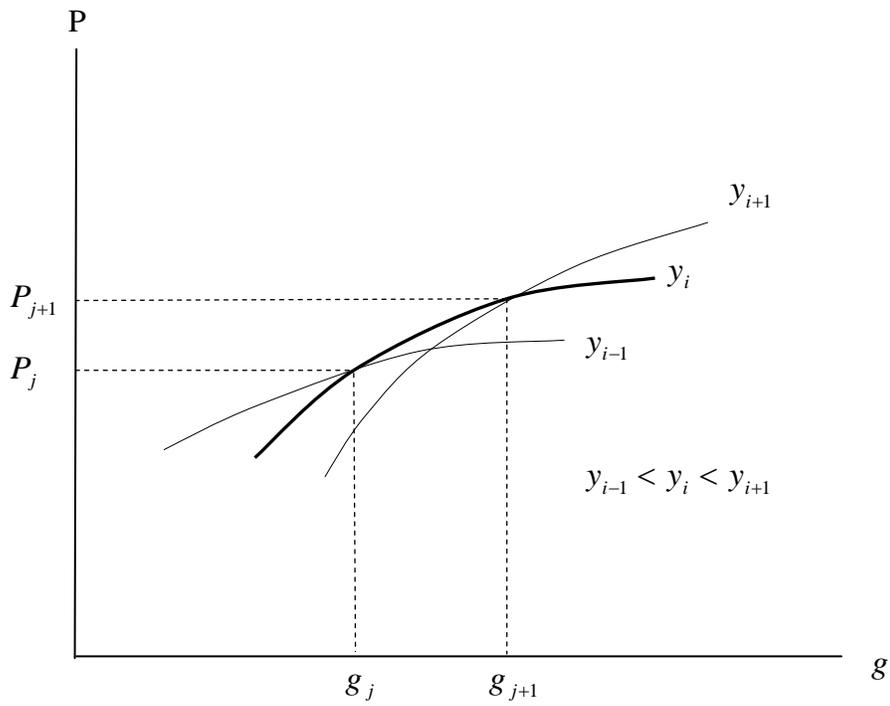


Figure 1. The Single Crossing Condition: Indifference Curves for Three Households

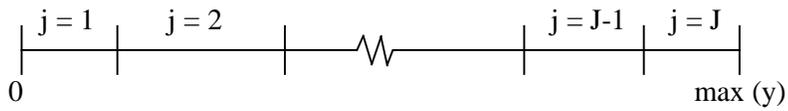


Figure 2. Partition of Households across Communities by Income

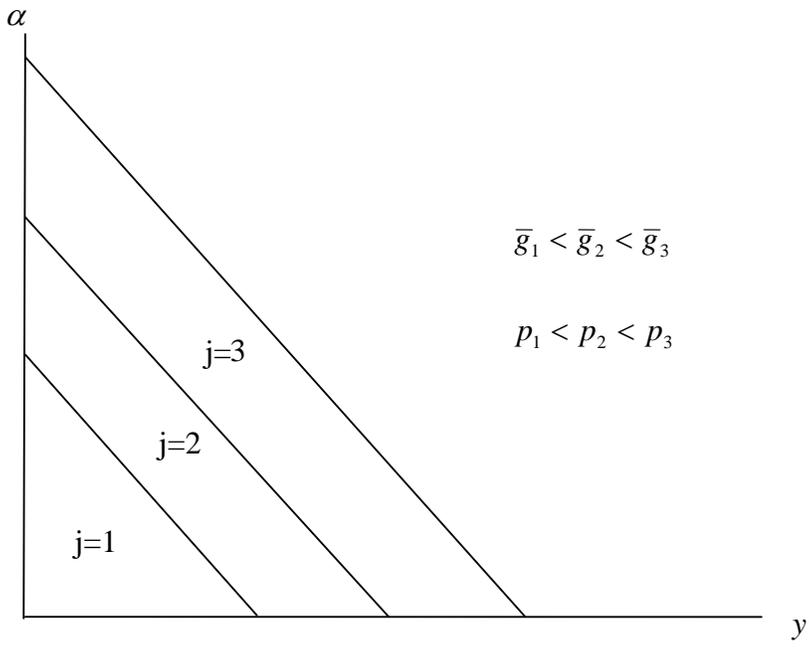


Figure 3. Partition of Households into Communities by Preferences and Income

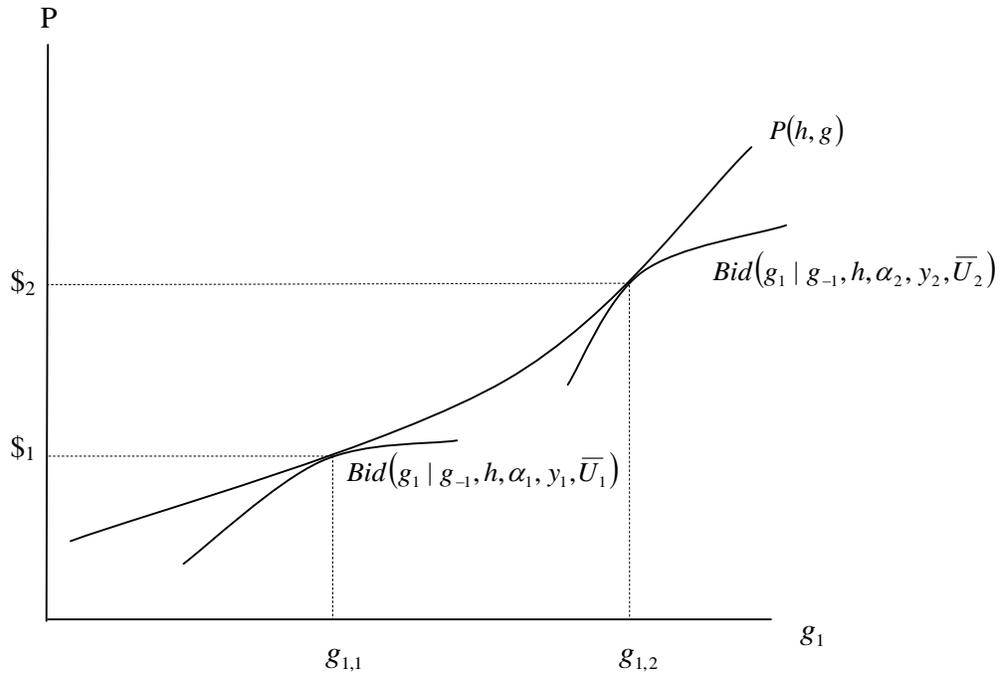


Figure 4. Bid Functions for Housing as a Function of g_1 in Hedonic Equilibrium

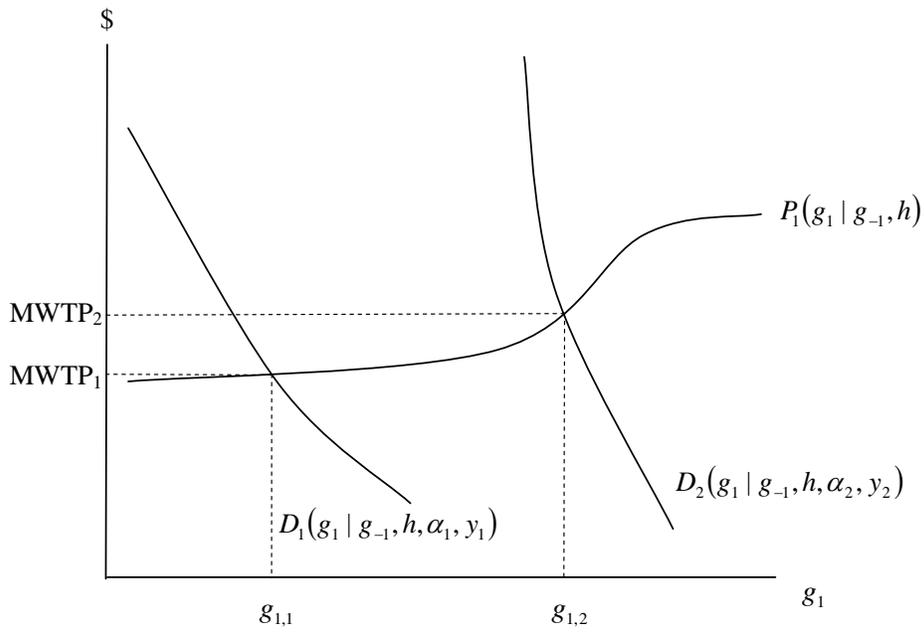


Figure 5. Implicit Price Function for g_1 and Demand Curves for Two Households