2009

Hedonic Markets and Explicit Demands: Bid-Function Envelopes for Public Services, Neighborhood Amenities, and Commuting Costs

John Yinger
Syracuse University, jyinger@maxwell.syr.edu

ISSN: 1525-3066

Center for Policy Research
Working Paper No. 114

Hedonic Markets and Explicit Demands: Bid-Function Envelopes for Public Services, Neighborhood Amenities, and Commuting Costs

John Yinger

Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
426 Eggers Hall
Syracuse, New York 13244-1020
(315) 443-3114 | Fax (315) 443-1081
e-mail: ctrpol@syr.edu

March 2009

$5.00

CENTER FOR POLICY RESEARCH – Spring 2009

Douglas A. Wolf, Interim Director
Gerald B. Cramer Professor of Aging Studies, Professor of Public Administration

Associate Directors

Margaret Austin
Associate Director
Budget and Administration

John Yinger
Professor of Economics and Public Administration
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Badi Baltagi .... Economics
Robert Bifulco………………… Public Administration
Kalena Cortes………………………………Education
Thomas Dennison .... Public Administration
William Duncombe .... Public Administration
Gary Engelhardt .... Economics
Deborah Freund .... Public Administration
Madonna Harrington Meyer .... Sociology
Christine Himes .... Sociology
William C. Horrace .... Economics
Duke Kao .... Economics
Eric Kingson .... Social Work
Sharon Kioko…………………..Public Administration
Thomas Kniesner .... Economics
Jeffrey Kubik .... Economics
Andrew London .... Sociology
Len Lopoo .... Public Administration
Amy Lutz………………………………… Sociology
Jerry Miner .... Economics
Jan Ondrich .... Economics
John Palmer .... Public Administration
David Popp .... Public Administration
Christopher Rohlfs .... Economics
Stuart Rosenthal .... Economics
Ross Rubenstein .... Public Administration
Perry Singleton…………………………Economics
Margaret Usdansky .... Sociology
Michael Wasylenko .... Economics
Jeffrey Weinstein…………………………Economics
Janet Wilmoth .... Sociology

GRADUATE ASSOCIATES

Sonali Ballal .... Public Administration
Jesse Bricker .... Economics
Maria Brown .... Social Science
Qianqian Cao .... Economics
Il Hwan Chung .... Public Administration
Qu Feng .... Economics
Katie Fitzpatrick .... Economics
Virgilio Galdo .... Economics
Jose Gallegos .... Economics
Julie Anna Golebiewski .... Economics
Nadia Greenhalgh-Stanley .... Economics
Clorise Harvey .... Public Administration
Becky Lafrancois .... Economics
Hee Seung Lee .... Public Administration
John Ligon .... Public Administration
Allison Marier .... Economics
Larry Miller .... Public Administration
Mukta Mukherjee .... Economics
Phuong Nguyen .... Public Administration
Wendy Parker .... Sociology
Kerri Raissian .... Public Administration
Shawn Rohlin .... Economics
Amanda Ross .... Economics
Jeff Thompson .... Economics
Tre Wentling .... Sociology
Coady Wing .... Public Administration
Ryan Yeung .... Public Administration

STAFF

Kelly Bogart.…….………….…Administrative Secretary
Martha Bonney……Publications/Events Coordinator
Karen Cimilluca…………………..Office Coordinator
Roseann DiMarzo.…….…… Administrative Secretary
Kitty Nasto……..……………Administrative Secretary
Candi Patterson......................Computer Consultant
Mary Santy……..……………Administrative Secretary

Abstract

Hedonic regressions with house value as the dependent variable are widely used to study the value of public services and amenities. This paper builds on the theory of household bidding and sorting to derive a bid function envelope, which provides a form for these regressions. This approach uses a general characterization of household heterogeneity, yields estimates of the price elasticities of demand for services and amenities directly from the hedonic with no need for a Rosen two-step procedure, and provides tests of key hypotheses about household sorting. An application to data from Cleveland in 2000 yields precise estimates of price elasticities for school quality, distance from environmental hazards, and neighborhood ethnic composition. The results support the sorting hypotheses and indicate that household preferences are very heterogeneous, with some households placing a negative value on many "amenities."

John Yinger is a Professor of Economics and Public Administration at the Maxwell School, Syracuse University. I am grateful to Professor David Brasington; this paper would not have been possible without the data he posted on the internet and he also gave me helpful comments. Additional helpful comments were provided by Steve Ross, and Phuong Nguyen helped me with the literature review.

JEL Codes: H73, R21

Key Words: Hedonics, capitalization, bidding, sorting

Introduction

House-value regressions, also called hedonic regressions, are one of the central empirical tools of urban economics and local public finance. This tool has been used to study a wide range of topics, including the willingness to pay for public services, the willingness to pay for environmental quality, the impact of property taxes on housing markets, the trade-off between housing and commuting costs, and racial prejudice and discrimination. Scholars have also produced an extensive conceptual literature on each of these topics. With a few important exceptions, however, the conceptual literature has not been used to determine how house-value regressions should be specified. The main purpose of this paper is to bring the conceptual and empirical literatures on these topics closer together by introducing the logic of bid functions directly into the specification of a house-value regression. This approach makes it possible to test some basic tenets of bid-function theory, facilitates consideration of household heterogeneity, and leads to direct estimation of amenity demand elasticities.

The foundation of this paper is the theory of household bids for housing in locations with different characteristics. The framework developed here brings together bids for housing based on public services, neighborhood amenities, commuting costs, and local taxes. This framework extends the literature in two key directions. First, I introduce constant-elasticity demand functions for public services (or neighborhood amenities) and housing, which have been widely used in studies of those topics. This step leads to bid functions that can be incorporated into house value regressions and then estimated. The functional form I derive can accommodate many different public services and amenities, as well as commuting costs and local taxes. Moreover, this derivation reveals that the service/amenity and housing price elasticities appear directly in bid functions and can be estimated with a one-step hedonic procedure.

Second, I introduce household heterogeneity into the public service, amenity, and housing demand functions. Using a very general characterization of this heterogeneity and the main theorem from the literature on household sorting (namely, that households sort according to the slopes of their bid functions), I derive the envelope of the household bid functions across household types and show how this envelope can be incorporated into a house-value regression. This derivation shows how some studies confound the amount a given type of household is willing to pay for an increment in public services (or in an amenity or in access to jobs) with movement along the bid-function envelope, which involves a change in household type. My approach shows how to separate these two effects and leads to a test of the sorting theorem. This approach also can accommodate cases in which an "amenity" has positive value for some households and negative value for others, and extension that proves to be important empirically.

The second part of the paper estimates this new house value regression using data for all the house sales in Cleveland in 2000. I find strong impacts on house values from school district characteristics, the crime rate, and distance from environmental hazards and support for the sorting theorem. The regressions also indicate that the share of a neighborhood that is black and the share that is Hispanic are disamenities for some households and amenities for others.

Preview of the Literature

This paper touches on many topics, including household bidding and sorting, hedonic equations, the capitalization of property taxes and public services into house values, the impact of air pollution on house values, the trade-off between commuting costs and housing prices, and racial prejudice and racial residential segregation. This section provides a brief introduction to the literatures on these topics. A more extensive discussion is provided throughout the paper as each topic arises in the development of my model and empirical strategy.

This paper begins with the theory of household bidding and sorting. Many scholars have contributed to a standard model that explains how households bid for housing across communities with various levels of public services and property taxes and how households sort into communities. Key contributions to this literature have been made by Ellickson (1971), Epple, Filimon, and Romer (1984, 1993), Epple and Platt (1998), Henderson (1977), and Wheaton (1993). This literature is reviewed in Ross and Yinger (1999). This theoretical literature ties to empirical work because it predicts that property taxes and public service levels will be capitalized into the price of housing.

A large empirical literature on capitalization, much of it inspired by Oates (1969), has also appeared. Recent contributions on public service capitalization include Black (1999), Brasington (2002, 2004, 2007); Brasington and Haurin (2006); Bogart and Cromwell (1997); Clapp, Nanda, and Ross (2008); Downes and Zabel (2002); Kane, Riegg, and Staiger (2006); and Weimer and Wolkoff (2001). Studies on tax capitalization are reviewed in Yinger et al. (1988) and Ross and Yinger (1999).

A related branch of the literature that is also central to this paper concerns hedonic equations, which are designed to estimate the impact of product attributes on product prices. In the case of housing, the product attributes include characteristics of the house's location, such as school quality. This literature, which is reviewed in Sheppard (1999) and Taylor (2008), goes beyond estimating the impact of public services on house values by estimating the underlying demand for the public services. Most studies follow the two-step procedure in Rosen (1974). The first step is to estimate a hedonic equation. The second step is to find implicit prices of product attributes by differentiating the hedonic equation and then to estimate the demand for each attribute as a function of its price.

Brown and Rosen (1982), Epple (1987), and others explain that the main challenge to this approach is that the hedonic equation is likely to be nonlinear so that the implicit price depends on the quantity of the attribute selected by the household and therefore is endogenous. Much of the literature has been devoted to solving this endogeneity problem. Epple (1987), Ekeland, Heckman, and Nesheim (2004), and Bajari and Kahn (2005) provide alternative estimating techniques. A striking feature of my approach is that because it estimates service/amenity price elasticities directly, benefit calculations for changes in public service or amenity levels can be estimated without a second step.

The concept of sorting is brought into the estimation of a hedonic equation by Epple and Sieg (1999), Epple, Romer, and Seig (2001), and Bayer, Ferreira, and McMillan (2007). The first two studies derive and estimate a general equilibrium model in which households sort into communities and select the level of local public services. Bayer, Ferriera, and McMillan address the household allocation problem using a discrete choice model in which households are allocated across houses according their valuation of housing and neighborhood attributes. A detailed comparison of my approach with these two approaches is provided later in this paper.

The capitalization and hedonics literatures have also been applied to topics other than local public services. Polinsky and Shavell (1976) extend a basic urban model to include an exogenous amenity such as air pollution, and many studies, such as Chay and Greenstone (2005), Brasington and Hite (2005), and Bui and Mayer (2003), estimate the impact of pollution on property values. Yinger (1976) extends a basic urban model to include an endogenous amenity, with a focus on neighborhood ethnic composition.[1] Empirical studies on this topic include Kiel and Zabel (1996), Bayer and McMillan (2008), and Zabel (2008).

Finally, this paper builds on the central result of the Alonso/Mills/Muth model of urban residential structure, namely, that commuting costs are reflected in housing prices. This model is reviewed in Brueckner (1988), and previous tests of the central result can be found in Yinger (1979) and Coulson (1991).

The Theory of Bidding and Sorting

This section provides a comprehensive bidding framework that encompasses public services, neighborhood amenities, commuting distance, property taxes, and local income taxes. This framework builds on the consensus bidding framework that has been developed in the literature, but gives it a new link to empirical application by introducing constant elasticity demand functions for public services and amenities. This section also shows how to incorporate this bidding framework into a house-value regression and develops a new method to account for household heterogeneity and sorting. It also shows how to extend this approach to many amenities, including air pollution and neighborhood ethnic composition, and to commuting costs, and it explains how this approach compares to others in the literature.

Housing Bids and Locational Equilibrium with Public Services

This standard model assumes that households maximize utility over public services, housing, and a composite good. Households make bids on housing based on public service levels and local property tax rates, and households with different incomes and preferences sort into different jurisdictions. This model assumes that households are mobile, so a key equilibrium condition is that all households in an income-taste class achieve the same utility level. Households are assumed to live in a metropolitan area with many local governments financed by a property tax. All the people who live in a given jurisdiction are assumed to receive the same level of public services, and the only way to gain access to the public services in a jurisdiction is to live there. A few scholars have examined local income taxes in this setting (Goodspeed 1989; Pogodzinski and Sjoquist 1993), and I consider both local property and income taxes. This paper also assumes that all households are homeowners. Depending on assumptions about property tax incidence, the models presented here could be applied to renters, as well.

In this model, a household maximizes a utility function over housing services, H; the quality of public services, S; and a numeraire good, Z, subject to a budget constraint. This constraint sets household income, Y, equal to Z plus housing consumption, PH, where P is the price per unit of H, plus property taxes. A household's property tax payment is the effective tax rate, τ, multiplied by its house value, $V = PH/r$, where r is a discount rate. In symbols,

$$Y = Z + PH + \tau V = Z + PH\left(1 + \frac{\tau}{r}\right) = Z + PH(1 + \tau^*) \quad (1)$$

where $\tau^* = \tau/r$.

A straightforward way to derive housing bids is to determine the maximum amount a household would pay per unit of H in different locations, holding utility constant.[2] Solving equation (1) for P, this approach leads to the following maximization problem:

$$\begin{aligned} &\text{Maximize} \quad P = \frac{Y - Z}{H(1 + \tau^*)} \\ &\text{Subject to} \quad U\{Z, H, S\} = U^0 \end{aligned} \quad (2)$$

where U^0 is the utility level obtained by households in this income-taste class. In this problem, Z and H are the choice variables, and S and τ are parameters. As a result, the envelope theorem can be used to determine the impact of S and τ on bids. Specifically,

$$P_S = \frac{U_S / U_Z}{H(1 + \tau^*)} = \frac{MB_S}{H(1 + \tau^*)} \quad (3)$$

and

$$P_t = -\frac{P}{(r+\tau)} = -\frac{rP}{(1+\tau^*)} \tag{4}$$

where subscripts indicate partial derivatives. The key to interpreting equation (3) is to recognize that its numerator is the marginal rate of substitution between *S* and the numeraire, which is simply the height of the demand curve, also called the marginal benefit from *S* or MB_S.

The differential equation (4) can be solved using the initial condition that the before-tax price, $\hat{P}$, which depends on *S*, equals the after-tax price, *P*, when τ equals zero. The solution is

$$P\{S,t\} = \frac{\hat{P}\{S\}}{(1+\tau^*)} \tag{5}$$

Differentiating (5) with respect to *S* yields another helpful result:

$$P_S = \frac{\hat{P}_S}{(1+\tau^*)} \tag{6}$$

Bringing in an Public Service Demand Function

Now let us part from the literature by assuming that the demand for public services takes a constant-elasticity form. This form has been widely used and works well in demand functions estimated with community-level data (Duncombe and Yinger 1998). This form is

$$S = K_S N^{\delta} Y^{\theta} W^{\mu} \tag{7}$$

where *W* equals tax price, K_S is a constant, and *N* is a vector of preferences and perhaps other variables that influence the demand for *S*. Now the inverse demand function for *S* can easily be derived by solving equation (7) for W:

$$W = \left(\frac{S}{K_S N^{\delta} Y^{\theta}}\right)^{1/\mu} \equiv MB_S \tag{8}$$

The constant-elasticity form has also been widely used for housing (Zabel 2004):

$$H = K_H M^{\rho} Y^{\gamma} \left(P(1+\tau^*)\right)^{\nu} \tag{9}$$

where K_H is a constant and M is a vector of preference and perhaps other variables that influence housing demand. Using equation (5), this equation can be re-written as

$$H = K_H M^{\rho} Y^{\gamma} \hat{P}^{\nu} \tag{10}$$

Combining equations (3), (6), (8), and (10) yields

$$\hat{P}_S = \frac{MB}{H} = \frac{\left(\dfrac{S}{K_S N^{\delta} Y^{\theta}}\right)^{1/\mu}}{K_H M^{\rho} Y^{\gamma} \hat{P}^{\nu}} = \frac{S^{1/\mu}}{\left(K_S N^{\delta}\right)^{1/\mu} K_H M^{\rho} Y^{(\theta/\mu)+\gamma} \hat{P}^{\nu}} \tag{11}$$

or

$$\hat{P}_S \hat{P}^{\nu} = \frac{S^{1/\mu}}{\left(K_S N^{\tau}\right)^{1/\mu} K_H M^{\rho} Y^{(\theta/\mu)+\gamma}} = \psi S^{1/\mu} \tag{12}$$

where

$$\psi = \left(\left(K_S N^{\tau}\right)^{1/\mu} K_H M^{\rho} Y^{(\theta/\mu)+\gamma}\right)^{-1} \tag{13}$$

Equation (12) is an exact differential equation, which can be solved by integrating both sides:

$$\left(\frac{1}{1+\nu}\right)\hat{P}^{1+\nu} = C + \left(\frac{\psi\mu}{1+\mu}\right) S^{(1+\mu)/\mu} \tag{14}$$

where C is a constant of integration. Solving for $\hat{P}$ yields

$$\hat{P}\{S\} = \left(\left(1+\nu\right)\left(C + \left(\frac{\psi\mu}{1+\mu}\right) S^{(1+\mu)/\mu}\right)\right)^{1/(1+\nu)} \tag{15}$$

Sorting

The next step is to allow for multiple types of households. According to the standard model, different household types sort into different jurisdictions because their bid functions have different slopes with respect to S. As shown by Ellickson (1971), Henderson (1977), Yinger (1982, 1995), Epple, Filimon, and Romer (1984, 1993), and Wheaton (1993), the household type with the steepest bid function wins the competition for housing in the jurisdiction with the highest-quality public services. Several scholars build models with continuous household distributions, not discrete income-taste classes. In these models, household types are determined by a distribution of income (Epple, Filimon, and Romer 1984, 1993) or of both income and preferences (Epple and Platt, 1998). This approach leads to the same general result: Households with steeper bid functions live in jurisdictions with higher S. A key assumption in all these models is the so-called single-crossing assumption, which says that if a household type has a steeper bid function than another at one value of S, it also has a steeper bid function at other values of S. My approach to sorting builds on the single-crossing assumption and the assumption that people sort into different jurisdictions according to the relative slopes of their bid functions.

The slope of household's bid function is the derivative of $\hat{P}$ with respect to S, which is given by equation (11). A household type's relative slope is determined by all the variables in equation (11) that are unique to that household, that is, everything except for S and $\hat{P}$. These terms are exactly the ones collected in ψ as defined by equation (13). In other words, ψ can be interpreted as in index of the steepness of a household's bid function—and hence and index of the value of S into which the household sorts.

To translate this observation into an estimating equation, we can derive the envelope of bid functions with various values of ψ. This envelope defines the set of winning bids. Consider two household types whose bid functions cross at a given value of S, say S*. These households have different values of ψ and hence different bid-function slopes, but, by the definition of "cross," they also have the same bid, $\hat{P}$, at S*. As shown in Figure 1, therefore, the bid function with the steeper slope must have a smaller intercept. The derivation of an envelope therefore involves finding a formula for the constant term such that $d\hat{P}/d\psi = 0$ when S is held constant at

$S\{\psi\}$, that is, at the value of S associated with the "winning slope". Applying this condition to equation (15), we find that

$$\left.\frac{dC}{d\psi}\right|_{S=S\{\psi\}} = \frac{-\mu}{1+\mu} S\{\psi\}^{\frac{1+\mu}{\mu}} \tag{16}$$

To solve this differential equation, that is, to find the envelope, we need to specify a form for $S\{\psi\}$. This form depends on the distribution of the factors that influence ψ, as well as on the distribution of S and the sizes of the jurisdictions with different values of S. In other words, it depends on the demand for and supply of S. I do not derive this form theoretically, but instead approximate it with a general form, the parameters of which can be estimated. The most general form I have been able to find that leads to a differential equation with an analytical solution is

$$S\{\psi\} = (\sigma_1 + \sigma_2\psi)^{\sigma_3} \tag{17}$$

Substituting equation (17) into equation (16) and integrating leads to:

$$C = C_0 - \left(\frac{\mu}{1+\mu}\right)\left(\frac{\mu}{\sigma_3(1+\mu)+\mu}\right)\left(\frac{(\sigma_1+\sigma_2\psi)^{\frac{\sigma_3(1+\mu)+\mu}{\mu}}}{\sigma_2}\right) \tag{18}$$

where C_0 is a constant of integration. Solving equation (17) for ψ and substituting the result and equation (18) into equation (15) leads to the bid envelope, identified with superscript E:

$$\hat{P}^E\{S\} = \left((\nu+1)\left(C_0 - \left(\frac{\sigma_1}{\sigma_2}\right)\left(\frac{\mu}{1+\mu}\right)S^{\left(\frac{1+\mu}{\mu}\right)} + \left(\frac{1}{\sigma_2}\right)\left(\frac{1}{\frac{1+\mu}{\mu}+\frac{1}{\sigma_3}}\right)S^{\left(\frac{1+\mu}{\mu}+\frac{1}{\sigma_3}\right)}\right)\right)^{1/(\nu+1)} \tag{19}$$

This envelope is illustrated in Figure 2 (using assumed values for the parameters). Equation (19) makes it clear that the impact of S on bids depends not only on household's willingness to pay for S, as indicated by μ, but also on the sorting process, as measured by the σ parameters. To put it another way, any estimated impact of S on housing prices describes movement along the bid-function envelope, not the willingness to pay of a particular type of household.[3]

Further explanation of this approach is provided by Figure 3, which is an extension of figures in Follain and Jimenez (1985) and Taylor (2008). The bid-function envelope, two illustrative bid functions for specific household types, and two illustrative offer functions for housing providers appear in the top panel. The slopes of these functions are plotted in the bottom panel. Even though people with steeper bid functions sort into higher-S locations, the slope of the envelope declines as the level of S increases because all the underlying individual bids are affected by movement down the individual demand curves for S. This decline in the slope of the envelope is moderated by the increase in the slope of the underlying bid functions as groups with steeper slopes win the competition at higher levels of S. An illustrative $S\{\psi\}$ function, which captures the second effect, appears in Figure 3. This function is anchored by bid function i; it indicates how much the slope must increase to go from the slope of bid function i (which includes just the first effect) to the slope of the envelope (which includes both effects).

The positive slope of the $S\{\psi\}$ function corresponds to the theorem that household types with steeper bid functions win the competition for housing at higher levels of S. It follows that we can test this theorem by looking at the sign of the estimated slope of this function, which is determined by the sign of σ_2 in equation (17). So long as σ_3 is positive, which is assumed in all my estimations, a positive sign for σ_2 supports the sorting theorem.

Generalizing to Many Services or Amenities

A key advantage of this approach is that, under fairly general assumptions, it can be applied to the case of many public services and to amenities, such as clean air. Suppose S is a vector of public service outcomes and amenities, indexed by i, each with its own constant-elasticity demand function. Then there is a separate version of the above equations for each S_i. If the demand for S_i does not depend on any S_j, $j \neq i$, then the expressions for S_j can be treated as part of the constant of integration, C, in equation (19) for S_i, and the envelope with an S vector is

$$\hat{P}^E\{S_1,...,S_n\} = \left((\nu+1)\left(C_0 + \sum_i \left(\left(\frac{1}{\sigma_{2i}\eta_i}\right)(S_i)^{\eta_i} - \left(\frac{\sigma_{1i}}{\sigma_{2i}\zeta_i}\right)(S_i)^{\zeta_i}\right)\right)\right)^{1/(\nu+1)} \tag{20}$$

where

$$\eta_i = \left(\frac{1+\mu_i}{\mu_i} + \frac{1}{\sigma_{3i}}\right) \quad \text{and} \quad \zeta_i = \left(\frac{1+\mu_i}{\mu_i}\right) \tag{21}$$

In standard demand theory, the first-order conditions for utility maximization are solved to yield a set of demand equations, and the demand for X_i is not a function of X_j, although it may be a function of the price of X_j. In the analysis considered here, however, a household is not selecting S_i and S_j, but is instead deciding how much to bid at given values of S_i and S_j. If S_i and S_j are complements or substitutes, therefore, the amount that a household bids to live in a community with a certain value of S_i may depend on how much S_j the community provides. If this dependence is direct, the derivation of the household-type-specific bid functions underlying equation (20) runs into a problem. To be specific, the derivative of this bid function with respect to S_j now has two terms, equation (11) and a new term representing the impact of a change in S_j on the household's willingness to pay for S_i. Differentiating equation (15) no longer leads to the same answer as integrating equation (11).

To resolve this problem, I assume that S_i may depend on the determinants of S_j, but not on S_j itself. If S_i and S_j are complements, for example, households may demand more S_i in a community that has relatively low costs for producing S_j, and is therefore more likely to select a high value of S_j. The systematic determinants of S_j can be treated as components of the N vector in equation (7) for S_i, and therefore cause no complications for the derivation. An equivalent assumption is that households demand for S_i reflects the systematic determinants of S_j but not the unobserved and random factors. It is as if the household expects the systematic factors to determine the level of S_j, at least in the long run, so they ignore the S_j error term when thinking about the impact of S_j on their bids associated with S_i.

In addition, an extension to many amenities must assume that the table of values for the various amenities is full enough, as a first approximation, so that each amenity can be selected independently of all others. A high correlation between two measures of school quality, for example, might force households to select from a limited number of amenity pairs, so that the marginal conditions derived here do not hold. In this case, it would be appropriate to combine these two measures into a single index and to treat this index as the amenity measure, which is the procedure I follow below in averaging passing rates across various state tests.

Neighborhood Ethnic Composition

Scholars have long recognized that a neighborhood's ethnic composition might influence house values for at least three reasons (Kiel and Zabel 1996). First, ethnic prejudice may lead some households to prefer neighborhoods in which their own ethnic group is concentrated. According to the survey evidence, some people in every ethnic group have this type of prejudice (see, for example, Charles 2000), so prejudice alone is unlikely to generate a monotonic relationship between a neighborhood's ethnic composition and housing price.[4]

Second, historically disadvantaged minority groups tend to cluster in neighborhoods with poor amenities, even after controlling for income (Deng, Ross, and Wachter 2003). This clustering presumably reflects past discrimination and the attendant disparities in wealth and other factors that influence housing choices. To the extent that these poor amenities cannot all be observed, ethnic composition might serve as a proxy for them. Not all unobserved amenities in largely minority neighborhoods need be negative, of course. People in a certain ethnic group might bid more to live in neighborhoods where their group is concentrated to gain access to ethnic restaurants or social organizations that are located there.

Third, households belonging to certain ethnic groups might continue to face discrimination in housing and mortgage markets. Although the extent of discrimination has declined over time, recent audit studies have found evidence of continuing discrimination against blacks and Hispanics in these markets (Ross and Turner 2005; Ross et al. 2008). Discrimination could affect our estimated hedonic regression in three ways. First, price discrimination against certain groups could lead to higher housing prices in neighborhoods where those groups are concentrated. This effect seems unlikely, however; recent studies find no evidence of price discrimination in housing markets (Ross and Turner 2005). Second, discrimination could push blacks and Hispanics into neighborhoods with poor amenities, thereby reinforcing the correlation between unobserved neighborhood traits and ethnicity discussed above. One study based on 1989 housing audit data (Ondrich, Ross, and Yinger 2003) found that even after controlling for housing price, real estate brokers steered blacks, but not whites, toward neighborhoods that had lower average house values or higher average house ages than the one that was original requested. Third, discrimination might restrict the ability of minority households to move out of largely minority neighborhoods and thereby artificially boost housing prices there.

An unusual feature of the bid functions derived earlier makes it possible for them to accommodate cases in which an increment in an "amenity" has a positive impact on bids for some households and a negative impact for others. According to equation (17), a household's marginal willingness to pay for an amenity, ψ, is negative if

$$(\sigma_1)^{\sigma_3} \equiv \sigma^* > S \tag{22}$$

As a result, any bid function that is tangent to the envelope at a value of S below σ^* has a negative slope. The regressions presented below provide an estimate of σ^* for each amenity and therefore indicate whether some households have a negative ψ for each amenity.

Because a majority of the population in my sample is white, I define two "amenities" as the inverse of the black or Hispanic concentration in a neighborhood. People who prefer a white, non-Hispanic neighborhood will place a positive value on these "amenities." If some households (of any ethnicity) value an increase in the black or Hispanic population, then the estimated values of σ^* will be greater than the minimum value of these two "amenity" variables in the sample. Unlike previous work on this topic (including my own), therefore, this approach does not confuse the estimated bid envelope with the willingness to pay of any particular group. Instead, the degree of heterogeneity in household preferences is revealed by the estimation.

My specification also identifies neighborhoods in or near a black ghetto (defined as a set of contiguous neighborhoods with a population that is at least 80 percent black). These two variables control for additional price effects that reflect unmeasured neighborhood quality variables in the ghetto, expectations about potential racial transition near the ghetto, or discriminatory restrictions on the ability of blacks to move outside the ghetto.

This specification obviously cannot fully separate all these factors. A negative relationship between housing prices and percentage minority could reflect either prejudicial attitudes on the part of whites or the effect of unmeasured neighborhood variables correlated with minority composition. A positive relationship between housing prices and percentage minority over some range could reflect prejudicial attitudes on the part of minorities or a preference among some households for amenities that cluster in high-minority neighborhoods.

Similarly, a finding that housing prices are lower either in or near the black ghetto than elsewhere could reflect either attitudes or unmeasured amenities.

Commuting Costs

In a standard urban model, $\hat{P}$ depends on distance from the CBD (or other worksite), labeled *u*, and per-mile commuting costs, *t*. Without operating costs, which are discussed more fully below, an equivalent approach is to introduce commuting time, *T*, valued at ωY per hour. Adding ωYT to the household budget constraint leads to the equilibrium condition, $P_T = -\omega Y/H$. [5] Combining this condition with the housing demand equation (12) and recognizing that income now must be expressed net of commuting costs yields,[6]

$$\hat{P}_T = \frac{-\omega Y}{H} = \frac{-\omega Y}{K_H M^{\rho}(Y-\omega YT)^{\gamma}\hat{P}^{\nu}} = \frac{-\omega Y^{1-\gamma}}{K_H M^{\rho}(1-\omega T)^{\gamma}\hat{P}^{\nu}} \tag{23}$$

Multiplying through by $\hat{P}^{\nu}$ results in another differential equation, the solution to which is

$$\hat{P}\{T\} = \left((\nu+1)\left(C' + \frac{Y^{1-\gamma}(1-\omega T)^{1-\gamma}}{K_H M^{\rho}(1-\gamma)}\right)\right)^{1/(\nu+1)} \tag{24}$$

where C' is a constant of integration.[7] Because this equation does not involve amenities or public services, it can be brought into equation (20) through the constant of integration.

The literature contains two approaches to household heterogeneity in a distance-based bid function. Beckman (1969) and Montesano (1972) derive bid functions assuming that the distribution of income across households takes the form of a Pareto distribution. Hartwick, Schweizer, Varaiya (1976) derive and Yinger (1979) estimates a model with distance-based bid functions for different discrete income classes. The technique developed earlier for public services provides an alternative approach, namely the derivation of the envelope of the distance-based bid functions across household types. This technique provides a more general foundation for empirical work than previous ones because it allows for a continuum of household types defined by both Y and other (not necessarily observed) traits.[8]

To derive the envelope, we can re-write equation (26) as

$$\hat{P}\{T\} = \left((\nu+1)\left(C' + \frac{\psi_T (1-\omega T)^{1-\gamma}}{(1-\gamma)} \right) \right)^{1/(\nu+1)} \tag{25}$$

where

$$\psi_T = \frac{Y^{1-\gamma}}{K_H M^{\rho}} \tag{26}$$

Household sorting by distance from a worksite depends on the slope of the bid function with respect to time, as given by equation (23), or the equivalent result for distance (Glaeser, Kahn, and Rappaport 2008). The terms in equation (26) include all the terms that determine the slope of a household type's bid function relative to that of other household types. As a result, integrating ψ_T out of equation (25) yields the envelope of bid functions across household types.

An estimating equation can be derived with a linear approximation for the ψ_T function.

$$T = \sigma_{T1} + \sigma_{T2}\psi_T \tag{27}$$

This form leads to

$$\begin{aligned} \hat{P}^E\{u\} = \Bigg((1+\nu) \quad \Bigg(C_0 - \left(\frac{\sigma_{T1}}{\sigma_{T2}(1-\gamma)} \right) (1-\omega T)^{1-\gamma} \\ + \left(\frac{1}{\sigma_{T2}(1-\gamma)} \right) T (1-\omega T)^{1-\gamma} + \left(\frac{1}{\omega \sigma_{T2}(2-\gamma)(1-\gamma)} \right) (1-\omega T)^{2-\gamma} \Bigg) \Bigg)^{1/(1+\nu)} \end{aligned} \tag{28}$$

This bid-function envelope is illustrated in Figure 4. As with amenity-based bid functions, this result shows that attempts to estimate bid-function parameters without considering household types (that is, without considering the σ parameters) and sorting may lead to misinterpretation.

The assumption of no operating costs requires further comment. This assumption is particularly plausible for transit trips, at least in Cleveland, because the fare does not depend on distance. Moreover, even with automobile travel, operating costs of travel are far below time costs. Using assumptions discussed below, operating costs are about 12.5 percent of total costs per mile. If commuting speed does not depend on income, the assumption of no operating costs implies that the income elasticity of per-mile transportation costs, say χ, equals 1.0. As is well known, so-called normal sorting, with higher-income households living at greater distances from worksites, arises if γ is greater than χ. Most find that $\gamma < 1$ (Zabel 2004), which implies that that normal sorting will not occur. As pointed out by Wheaton (1977) and Glaeser, Kahn, and Rappaport (2008), however, the estimated elasticities may not lead to normal sorting even with operating costs. In fact, if operating costs equal 12.5 percent of total cost, χ equals (1-.125) = 0.875, which is still above most estimates of γ.

The assumption that operating costs equal zero does not imply that normal sorting cannot occur, however. Glaeser, Kahn, and Rappaport show that normal sorting can arise if mode-choice decisions are linked to income. Although the derivations presented here cannot account for mode choice, they can account for the relationship between commuting speed and income. A positive impact of income on commuting speed lowers χ and makes normal sorting more likely. Normal sorting can also arise from a correlation between Y and the M variables in the demand function for H. The standard derivation assumes that nothing else in the housing demand function is correlated with Y. If ρ, the correlation between M and Y, is positive however, then sorting could lead to a positive correlation between Y and u, which is the definition of normal sorting, even if $\gamma < 1$. To be specific, normal sorting arises if the total derivative of H with respect to Y is greater than χ, that is, if

$$\sum_i \rho_i \left(\frac{\partial M_i}{\partial Y} \frac{Y}{M_i} \right) + \gamma > \chi \tag{29}$$

Local Income Taxes

In Ohio, local school districts are allowed to levy income taxes. The income tax rates districts select lead to another dimension of bidding and sorting. With a local income tax imposed at rate y, a household's net income is $Y(1\text{-}y)$.[9] This leads to a bid function over values of y that takes exactly the same form as the bid functions over values of T, except that $(1\text{-}y)$ replaces $(1\text{-}\omega T)$. The bid function envelopes also take the same forms with this substitution (and, obviously, with σ_{y1} replacing σ_{T1} and σ_{y2} replacing σ_{T2}).

The Final Estimating Equation

Finally, following equation (1) we can combine bids with τ and H, which is a function of structural housing characteristics, X. Following the literature (see Ross and Yinger 1999), I also introduce a parameter, β, to indicate the degree of property tax capitalization. The result:

$$V = \frac{P^E H\{X\}}{r} = \frac{\hat{P}^E\{S_1, S_2, ..., S_n, u, y\} H\{X\}}{(r + \beta\tau)} \tag{30}$$

where $\hat{P}^E$ now contains the elements of equations (20) and (28), as well as the terms from the comparable equation for y. When this equation is estimated, the error term reflects unobserved housing characteristics and individual bids that differ from market bids because of the bargaining skill (or lack thereof) of the people involved. In addition, I specify $H\{X\}$ as multiplicative; Sieg et al. (2002) show that this form is consistent with the assumption used in this paper that V is the product of P (a function of locational characteristics) and H (a function of the Xs).

Recovering the Other Demand Parameters

Successful estimation of equation (20) and (28) (through equation (30)) yields the price elasticity of demand for each amenity and the income and price elasticities of demand for housing. Moreover, the estimation of the σ parameters makes it possible to solve equation (17) for ψ and hence to estimate equation (13) for each amenity. Because S does not appear on the right side of equation (13), this estimation avoids the endogeneity problem that arises with the Rosen two-step method.[10] One disadvantage of this approach is that, except for the case of Y, it cannot separate the impacts on S and H of a variable that appears in both demand functions.

The coefficient of the estimated income term for equation (26) provides a test of the standard condition for normal sorting around a worksite. In the case of an amenity, normal sorting is defined as a sorting equilibrium in which higher-income households live in locations with better amenities. As first shown by Henderson (1977), normal sorting based on an amenity requires that $\theta/\mu + \gamma < 0$, so the coefficient of the income term in equation (13) provides a test of this condition.[11] The standard formula for omitted variable bias implies that this test can be extended to consider the indirect effects of income, as in equation (29), by estimating this equation with income as the only explanatory variable.

Comparison with Previous Approaches

The first step in the Rosen (1974) approach is to estimate an approximation to the hedonic equation. This approach brings no theoretical guidance to the specification of this equation, so many scholars have used flexible estimating techniques. Several early studies, including Halvorsen and Pollakowski (1981), estimated this equation using the Box-Cox specification. Some more recent studies, including Ekeland, Heckman, and Nesheim (2004), estimate the price function using parametric or semi-parametric techniques. This paper differs from most previous research in that it derives a theoretical form for the hedonic equation, which, curiously enough, is similar to the Box-Cox form.[12]

In addition to functional form, studies of the first-step hedonic have addressed problems of endogeneity and omitted variables.[13] Brasington and Hite (2008) argue that housing prices link housing characteristics with household characteristics and thereby create an endogeneity problem. This problem appears in equation (15), which contains household characteristics, but disappears from equation (19) because the household characteristics have been integrated out.

Many scholars argue that omitted variables lead to biased estimates for service/amenity variables, and a variety of different solutions to this problem have been proposed. One well-known approach based on school-attendance zone boundaries was introduced by Black (1999) and appears in Kane, Riegg, and Staiger (2006) and Bayer, Ferriera, and McMillan (2007). Scholars using this approach create a series of fixed effects for houses within a certain distance of school attendance zone boundaries. These scholars argue that houses near the same boundary share many unobserved neighborhood traits; the fixed effects are intended to control for these traits and to eliminate any bias that would arise if they were left out of the regression.[14] A second approach is to take advantage of unusual boundary arrangements, such as those created when a municipality contains more than one school district, to control for non-school factors. Examples include Bogart and Cromwell (1997) and Weimer and Wolkoff (2001).

A third strategy is to estimate the hedonic with an instrumental variables technique to eliminate the omitted variable bias. Good instruments for this approach are difficult to find. Policy choices by higher levels of government lead to compelling instruments in some studies, such as the Chay and Greenstone (2005) study of air pollution or the Gibbons and Machin (2003) study of public schools. In most other cases, such as Weimer and Wolkoff (2001) or Downes and Zabel (2002), the instruments are selected from a list of neighborhood characteristics, and it is difficult to see why they should be considered exogenous.

The second step in the Rosen approach is to calculate implicit prices, which are the derivatives of the estimated hedonic with respect to each service or amenity [$\hat{P}_S^E H/(r+\beta t)$], and then to estimate a demand model for that service or amenity. Because households set the marginal benefit from an amenity equal to its implicit price, this implicit price can be used as the dependent variable in an estimation of an inverse demand function for *S*, such as equation (8), or as an explanatory variable in a direct demand function, such as equation (7).

As noted earlier, the principal challenge facing this step is endogeneity; with a nonlinear price function a household's unobserved traits may affect both quantity and price, that is, both *S* and *W* (the implicit price) in equations (7) and (8). Many scholars have developed ways to address this endogeneity problem. See, for example, Ekeland, Heckman, and Nesheim (2004). Bajari and Kahn (2005) assume a unitary price elasticity of demand for the *S*. In this case, multiplying equation (8) by *S* leads to a new regression with the implicit price multiplied by *S* as the dependent variable and only the exogenous variables in the denominator as explanatory variables. This approach cannot be generalized, however, to non-unitary price elasticities.

As explained earlier, the approach proposed here avoids this endogeneity problem in a different way. First, this approach makes it possible in principle to estimate μ directly, with no need for a two-step procedure. Second, the other demand parameters for *S* and *H*, can be estimated via equation (13), which does not have *S* on the right side for any value of μ.

Several recent studies have explicitly considered sorting in hedonic regressions.[15] One approach with similarities to the one used here can be found in Epple and Sieg (1999) and Epple, Romer, and Sieg (2001). These scholars estimate general equilibrium models in which (as in Epple and Platt 1998) people sort according to income and to a parameter in their utility function.[16] Their analysis includes the analog to a bid-function envelope for a single amenity and also avoids the standard endogeneity problem. This approach is based on an explicit form for the indirect utility function that is consistent with a constant elasticity demand function for *H* but not for *S*. These models are estimated using a complicated statistical procedure that solves for the parameter values that produce the best correspondence between the distributions of outcomes generated by the model and the distributions of outcomes in a sample of communities.

The great strengths of these studies are that they incorporate both household sorting and the determination of *S* through voting once sorting has taken place and that they explicitly solve for an equilibrium and incorporate the solution into the estimating method. My approach cannot do either of these things. Instead, my approach overcomes some of the empirical weaknesses of these studies, namely, that they are limited to aggregate (community level) data, that they are difficult to extend beyond a single amenity, that they consider only two dimensions of household heterogeneity, and that they require a complex statistical technique.

Another clever approach is provided by Bayer, Ferreira, and McMillan (2007).[17] They use a complex procedure to estimate a multinomial discrete choice model in which households have different preferences and each household is allocated to the house that gives it the highest utility (net of price). An equilibrium allocation is therefore built into their estimation procedure. By focusing on household choice, instead of household bidding, this approach also can directly estimate variation in the amount different types of households are willing to pay for each amenity.[18] The main disadvantages of this approach are that it is based on a highly restrictive linear utility function and that it requires complex estimating procedures.

Bayer, Ferreira, and McMillan (2007) also estimate a standard hedonic equation to obtain average preferences (average MWTP) for each amenity. A simplified version of their utility function for household *i* in house *h* is $V_{ih} = \alpha_{0i} + \sum_j \alpha_{ji} X_{hj} - \alpha_{hi} P_h$, where *V* stands for utility, the

X's denote housing characteristics, and P is the price of housing.[19] A bid function for a given household (or household type) is found by solving for the price, say P^*, that holds utility constant (at V^*) over housing and neighborhood traits. As a result, the bid function in this case is $P_{hi}^* = \left(\alpha_{0i} - V_i^*\right)/\alpha_{hi} + \sum_j \left(\alpha_{ji}/\alpha_{hi}\right) X_{hj}$. Households sort according to the steepness of this bid function, so households with higher values of $\left(\alpha_{ji}/\alpha_{hi}\right)$ live in locations with higher values of X_j. This sorting problem leads to a bid-function envelope, which is the market price of housing.

Consider a single amenity, simplify the bid function to $P_i^* = a_i + b_i X$, and assume that the sorting equilibrium can be described by $X_i = \sigma_1 + \sigma_2 b_i$. Then, following the procedure presented earlier, the bid-function envelope is $P^E = C - (\sigma_1/\sigma_2)X + X^2/(2\sigma_2)$, where C is a constant. This derivation can be extended to many X's, each of which will enter in quadratic form. Even with a linear utility function, in other words, the bid-function envelope (i.e. the hedonic equation) is a quadratic function of the X's.[20] By Bayer, Ferreira, and McMillan's own assumptions, therefore, the linear hedonic equation they estimate is mis-specified, and their argument that the estimated coefficients indicate average preferences is implausible.[21]

Epple and Sieg (1999), Epple, Romer and Sieg (2001), Bayer, Ferreira, and McMillan (2007) and Bajari and Kahn (2005) explicitly allow for heterogeneity in household preferences.[22] The approach developed here also accounts for observed and unobserved heterogeneity in the demand for S and H both in the first step (where it is integrated out) and through the error term in the second-step regression, namely, equation (13). The estimating equation (30) also contains H directly, and accounts for observed or unobserved determinants of demand through the impact of these factors on housing characteristics, X. Neither my approach, nor any other, can rule out the possibility of bias from omitted *X*'s, but this type of bias is unlikely to be significant with a data set in which many *X*'s appear, such as the Cleveland data set used in this paper.

Estimation Procedures

Equation (30) provides a one-step conceptual method for obtaining service/amenity demand elasticities. Although this equation could be estimated directly, we can ensure both that the coefficients of the structural housing characteristics are not biased because of their correlation with unobserved neighborhood characteristics and simultaneously that the locational dimension of housing prices is accurately measured by estimating it in two stages. The first stage is a standard hedonic regression of the log of sales price on housing characteristics (*X*), within neighborhood differences, and neighborhood fixed effects. The second stage uses the coefficients of the fixed effects as the dependent variable to determine how locational characteristics affect the bid envelope (and underlying bids) using the forms derived earlier, after controlling for an extensive set of geographic variables.[23]

Data

The data used for this study build on the data set described in Brasington (2007) and Brasington and Haurin (2006). This data set consists of all the house sales in Ohio in 2000; this study makes use of the 22,857-observation sub-sample for Cleveland. As discussed more fully below, the data set contains information on sales price, housing characteristics, housing location, census block group characteristics, school performance, and air quality, among other things. These data were supplemented with a variety of zip-code and neighborhood characteristics.

The Hedonic Regression with Geographic Fixed Effects

The first-stage hedonic faces two main challenges. The first is to define the fixed effects. The second is to account for locational characteristics that vary at the level of an individual observation. The Brasington data set provides extensive information at the census-block-group (CBG) level. A CBG is a small unit of geography, so it is ideal for using as a neighborhood. In some cases, however, a school district or cuts across a CBG. As a result, I break up CBG's that fall into more than one school district. My results are based on CBG/school units (still called CBGs for conciseness) with at least two observed house sales.

Six locational variables vary within a CBG: distance to worksites, to the nearest public elementary school, to the nearest private school, to the nearest environmental hazard, to Lake Erie, to the nearest black ghetto, and to the Cleveland airport. Because the hedonic equation includes CBG fixed effects, these variables are specified to measure only within-CBG variation. For the last five variables, this variation is defined by the difference between each distance for the individual house and the distance for the house's CBG. This approach is equivalent to including the exponential of these distances, multiplied by a parameter, in equation (31).

The specification of the worksite variables comes from standard urban theory, but to preserve the simplicity of the first-stage hedonic, the derivations rely on simpler forms than the ones used above. The first step is to identify worksites. After plotting (by latitude and longitude) zip-codes containing at least 8,000 jobs (as measured by the number of employees),[24] I identified five clusters of zip codes with at least 20,000 jobs and defined them as Cleveland's worksites. These worksites contained 74.5 percent of the jobs in the Cleveland metropolitan area. Figure 5 presents a map of job and worksite locations. Finally, I specified the center of each worksite as the employment-weighted centroid of the zip codes in the relevant set.

The second step is to assign houses (and CBGs) to worksites. My approach is to use standard (but simplified) urban theory to express bid functions around each work site and then, following Wieand (1987) and Yinger (1992), to define the boundary between two worksites as the set of locations at which the bid functions for the two worksites meet. With a Cobb-Douglas utility function, the bid function for worksite i can be written:

$$\hat{P}\{u_i\} = \left(\frac{k}{U_i^*} \right)^{1/\alpha} (Y_i - t_i u_i)^{1/\alpha} \tag{31}$$

where α is the share of income (after commuting costs) spent on housing, k is a constant that depends on α, U_i^* is the utility level achieved by people working at worksite i, and u, Y, and t are distance, income, and per-mile commuting cost. All households at a given worksite are assumed to be identical. Each household in the sample is assigned to the worksite associated with the highest bid at that household's location.

The value of u_i in equation (31) is the distance from a house in the sample to the center of a worksite based on latitude and longitude.[25] The value of u_i for a CBG is set equal to the distance from the center of that CBG to the center of the worksite assigned to the largest number of houses in the sample in that CBG.[26] To make these worksite assignments, Y_i is set equal to annual payroll per employee for worksite i (an employment-weighted average over the relevant zip-codes) divided by 220 (to turn it into a daily amount). Following Lipman (2006), α is set at 25 percent (rounded from her estimate for 2000 of 27.5 percent for households in 28 metropolitan areas). In addition, I assume that one-way commuting cost per mile (half of t_i) equals operating cost per mile plus the value of commuting time (as a share of the hourly wage) multiplied by the hourly wage ($Y_i/8$) and divided by commuting speed. According to the U.S, Department of Transportation (2008), operating cost in 2000 was 12.2 cents per mile. Following Small, Winston, and Yan (2005), the value of commuting time is set at 90 percent of the wage rate (rounded from their estimate of 93 percent). Following Glaeser, Kahn, and Rappaport (2008), commuting speed is set at 37.5 MPH.

All that remains is to select the value of U_i^*. Raising U_i^* lowers the bids for worksite i and therefore lowers the number of houses assigned to that worksite. My approach is to vary the values of U_i^* for the five worksites until the share of employed people assigned to each worksite equals the share of jobs at that worksite based on the Zip-code data.[27] The number of employed people assigned to a worksite is the sum of the employed people in each of the CBGs assigned to that worksite. I do not observe employment directly. Instead, I approximate employment in each CBG as total population multiplied by the share of people between ages 18 and 64 years old, multiplied by one minus the unemployment rate. These employment calculations are based on all households in a CBG, not just on the households in the sample.

Once a worksite has been selected for each house in the sample, the first term of equation (31) is not relevant for the hedonic regression. Since the first term does not vary within a CBG (indeed, it does not vary for an entire worksite), it drops out of the hedonic with CBG fixed effects, and all that appears is the second term—with a value of u that is determined by its assignment to a worksite. The end product of all this work, therefore, is a series of 5 variables, each defined by the second term in the above equation, indexed by worksite: $(Y_i\text{-}t_i\, u_i)$. To keep all variation associated with a CBG in the estimated fixed effects, this term is expressed relative to the comparable term for the CBG, which is $(Y_i\text{-}t_i\, u_i^*)$, where u_i^* is the distance to the worksite from the center of the CBG. Because the hedonic is specified in multiplicative form, these terms are logged and their coefficients can be interpreted as $(1/\alpha_i)$.

The definitions of the variables for the basic hedonic are presented in Table 1, and the results are presented in Table 2. The hedonic is estimated with the "areg" option in STATA based on 22,857 observations. The R-squared is 0.7899. The housing characteristics are, with the exception of "One Story," highly significant with the expected sign. The set of 1,685 geographic fixed effects is also highly significant (p=0.000). Of particular note are the commuting variables, which all are significant at the 0.1 percent level. Their coefficients are all close to 3.0, which has the reasonable implication that α, the share of income spending on H, is about one-third. Additional significant impacts on housing prices come from the distance from an environmental hazard (positive) and from Lake Erie (negative)..

The Housing Bid Envelope

The second equation is the housing bid envelope. The dependent variable is the set of coefficients for the CBG fixed effects. Hence, the number of observations, 1,685, is the number of CBGs with at least two observations.

The first step in estimating the bid envelope is to find the commuting time between each CBG and its assigned worksite. The Brasington data include the average commuting time in the CBG of full time workers who do not work at home. The procedures given earlier lead to a distance from the center of the CBG to the assigned worksite. My approach is to use regression to connect these two and then to predict commuting time based solely on distance to worksite. This step is important: As shown by Yinger (1993), using straight-line distance can lead to large measurement errors if locations that are equidistant from a worksite in terms of straight-line distance are not equidistance along actual roads.[28] Measurement errors can also arise if commuting speeds are not the same in all directions.

This analysis begins with dummy variables for each quadrant around each worksite (northeast, southeast, and so on). The use of quadrants fits Cleveland, which has transit and bus lines running to the northeast, southeast, and southwest out of downtown.[29] Then I regress commuting time on distance to the assigned worksite with a separate coefficient for each quadrant and with several control variables to account for commuting to nearby jobs. This regression is used to predict commuting time based on distance from worksite alone, which is T in equation (28).[30] The implicit commuting speeds that result from this procedure are correlated with income and the income elasticity of per-mile commuting costs is 0.8436, which implies that normal sorting could occur even with $\gamma < 1$.

The second step is to select service and "amenity" variables. As indicated in Table 3, I selected six of these variables: school district test scores, a high school quality indicator, safety, distance to the nearest environmental hazard, and two measures of ethnic composition. The test score variable measures to the ninth-grade passing rate on the five examinations (mathematics, reading, writing, science, and citizenship) that form the basis of the state's accountability system.[31] The high school variable is the ratio of the share of students receiving a high-pass on state tests in grade 12 to the drop-out rate, which is defined as drop-outs in a year relative to the district's enrollment in grades 7-12. This measure focuses on the extremes, that is, on the extent to which a district produces many top students (the numerator) or fails to serve many poor students (the denominator). These two school measures are both averaged over the two school years preceding 2000, namely, 1997-98 and 1998-99. The safety measure is one minus the crime rate, which equals total crimes per thousand people. The ethnic composition measures are the inverse of the percent black and of the percent Hispanic in the CBG.

Table 3 also lists the tax variables. I observe the school income tax rate, but this variable equals zero for most school districts and exhibits little variation elsewhere. As a result, I do not attempt to estimate the envelope for this variable but instead simply include it in the regression as a control. Ohio requires assessment at market value so the nominal property rates are roughly equivalent to effective rates.[32] Most school districts coincide with cities. I observe the school tax rate, the gross city tax rate (which includes school taxes), and the net city tax rate (which reflects exemptions from property taxes). These data lead to four property-tax variables: the school tax rate, the tax rate for non-school city services, the reduction in the tax rate due to exemptions, and an indicator for non-city districts, about 15 percent of the sample, for which the latter two rates are not observed.

The third step is to define a series of neighborhood control variables. These variables could also be considered amenities, but they are all either dummy variables or interactions with dummy variables so they have a value of zero for many CBGs and are therefore unsuitable for the functional forms derived earlier.[33] As shown in Table 4, these variables include worksite and county fixed effects, as well as indicators of location near a wide range of factors that might influence housing prices.[34]

A few variables require special comment. The eastern suburbs of Cleveland receive unusual amounts of snow. The snow-belt variables are designed to determine the extent to which prices are different in locations with particularly heavy snowfall. Snowfall maps from the years just preceding 2000 indicate that the heaviest snows are east of the town of Pepper Pike and about 10 miles from Lake Erie.[35] The quadratic specification of these variables estimates the distance from the lake at which the peak price in these eastern suburbs appears. The ghetto variables were defined by first plotting all CBGs with at least 80 percent black and then identifying large clusters of these CBGs. Two such clusters were found and the population-weighted latitude and longitude of their centers were calculated. A similar approach was applied to air pollution. I identified CBGs where unusually high amounts of pollutants were released into the air. The vast majority of these locations were clustered in three locations. I found the pollution-weighted centers of these three clusters and measured distance for each CBG to the nearest cluster. The final variables measure whether a CBG was less than 20 miles to the nearest cluster and the distance from the cluster if it was within 20 miles. These air-pollution variables were interacted with direction (northeast, southeast, and so on) to account for the possibility that air pollution is affected by wind direction.

These steps make it possible to estimate the full bid function envelope, equation (30) without $H\{X\}$. This equation is difficult to estimate in its full non-linear glory, so I assume that $v = -1.0$. With this assumption, the left sides of equations (20) and (28) are replaced by $\log\{V\}$ and the $(1+v)$ terms disappear from the right side. This assumption also makes it possible to estimate a coefficient for the property tax variables using the approximation that $\ln\{r+\beta\tau\} \approx \ln\{r\} + (\beta/r)\tau$. Second, I assume that $\gamma = 1.0$. This assumption leads to simplified versions of the commuting time variables. These assumed housing elasticities are somewhat larger in absolute value than most estimates (Zabel 2004); a later section provides rough checks on their validity.

The model was estimated with nonlinear least squares (using the STATA "nl" procedure with the "hc3" option to obtain robust standard errors). Starting values were found through a grid-search procedure to find the lowest sum of squared errors (SSE) for various values of the six price elasticities, one for each service or "amenity." For the safety variable, the lowest SSE was consistently found to be at -1, which calls for a different functional form. As a result, my final estimates assume that this elasticity is -1 and make no attempt to estimate its standard error.[36]

I estimated the model with different assumed integer values for σ_3, which influences the quality of the approximation for the $\psi\{S\}$ function. Raising the value of σ_3 appears to improve the approximation slightly, as indicated by the SSE, but it also raises the correlation between the two variables for each amenity and therefore makes estimation more difficult. The model starts to break down (as indicated by an inability to estimate a standard error for the constant term) when σ_3 reaches 6 and breaks down entirely (an inability to estimate most standard errors) when

σ_3 reaches 10. Somewhat surprisingly, however, the results are similar for different values of σ_3, at least when the model can be estimated.

Empirical Results

The main results are presented in Table 5 (service, amenity, tax, and commuting variables) and Table 6 (geographic controls). The first column of Table 5 presents results for a standard specification in which the service/amenity variables are entered linearly. All of these variables except the safety variable are statistically significant. In the second column, which sets all the price elasticities to -1, both terms are significant for all the service/amenity variables except the test-score variable. The remaining three columns estimate the full model with various values for σ_3. All of the estimated elasticities increase slightly in absolute value as σ_3 goes up.

Consider the fourth column, which is based on a quadratic approximation to the $\psi\{S\}$ function (σ_3=2). All of the service/amenity elasticities (except the one for safety which, as indicated earlier is assumed to be -1) are highly significant and less than 1.0 in absolute value. The elasticity is -0.134 for Test Scores, -0.623 for High School, -0.320 for Distance to Hazard and -0.358 for both ethnic composition variables. Moreover, the estimated values for the σ_1 coefficients are statistically significant in every case. As discussed earlier, the signs of the σ_2 coefficients provide tests of the hypothesis that sorting is based on the slope of the underlying bid functions. A positive sign supports this hypothesis, whereas a negative sign rejects it. The estimated values for the σ_2 coefficients all have a positive sign, but they are significant at the 5 percent level only for High School and Non Hispanic. These results therefore support the sorting hypothesis in these two cases but do not reject it for any case.

Most of the other variables in Table 5 perform as expected. With the full specification, both of the commuting time variables are statistically significant, as are two of the three property tax variables. The coefficients of the property tax variables equal the degree of capitalization (β) divided by the discount rate (r). With r = 3 percent, the significant results in the last three columns indicate a fairly low β of 8 to 10 percent.[37] The insignificance of the school tax rate is surprising and may indicate a correlation between this rate and unmeasured school quality. The weak significance of the income tax rate is not so surprising given its limited variation.

Figures 6 and 7 plot the estimated envelopes and illustrative underlying bid functions (equation (20)).[38] Figure 6 presents these functions for the school variables, Safety, and Distance from Hazard. The first and fourth panels (for Test Scores and Distance to Hazard) have similar shapes; they indicate that the neighborhoods with the lowest values of these variables have much lower prices than other neighborhoods, but that the price impact of these variables is fairly small once a moderate level of the service/amenity is reached. For Test Scores, moving from the minimum value (17.2) to the mean value (55.0) raises the envelope by 252.1 percent and the illustrated bid function by four times as much. In contrast, moving from the mean to the maximum (91.0) raises the envelope by only 1.18 percent and hardly affects the illustrated bid. The effects of Distance to Hazard are much smaller. Moving from a very low value (0.1 miles) to the mean value (0.98 miles) raises the envelope by 14.5 percent, but moving from the mean to the maximum (4.9 miles) raises it by only 0.8 percent. One curious feature of both these figures is that the envelopes (and hence the underlying bid functions) have a negative slope for the lowest observed values of the service or amenity. This does not appear to be a substantive result, however, as it applies to only 30 observations in the case of Test Scores and only 10 observations in the case of Distance to Hazard.

The second panel reveals a different pattern for High School. Raising the high school variable from its median value (3.54, which is about the minimum point on the envelope) to its maximum (177.2) raises the bid envelope by 26.9 percent and would raise bids on the illustrated bid function by 43.2 percent. Below the median value of this variable, however, the envelope has a negative slope. Going from the minimum to the median value would lower the envelope by 37.5 percent. This result demonstrates the importance of accounting for household heterogeneity. In about half of the neighborhoods in this sample, improved high school quality, as measured by High School, leads to higher housing prices. These results also imply, however, that the people who live in school districts with the poorest high school quality would rather see their schools provide other types of outcomes (holding tax rates constant). Because High School focuses on the extremes, one possibility is that parents in these schools would rather see a concentration on students in the middle of the distribution that on students at the top and bottom. I explore these issues further in the following section.

The third panel of Figure 6 shows that Safety also has a positive impact on house values except at the bottom of the Safety distribution. Going from a value of 0.894 (the minimum point on the envelope and the 25th percentile of the Safety distribution) to the maximum value (0.999), raises the envelope by 9.9 percent. As shown in the figure, increased safety has a small negative impact on the envelope below a value of 0.894.[39] As in the case of High School, these results demonstrate the extent of heterogeneity in household preferences. Apparently some households would like their city to shift resources away from safety toward other (unspecified) things.

Figure 7 presents results for the ethnicity variables. The first panel shows that the bid envelope declines with percent black up to a neighborhood composition of 75 percent black. More specifically, going from zero to 75 percent black lowers the envelope by 27.1 percent, whereas going from 75 to 100 percent black raises the envelope by 2.9 percent. These results support the view that households have heterogeneous preferences for black neighbors and are illustrated by the two bid functions in this figure, one for a household type that prefers white neighborhoods and the other for a household type that prefers black neighborhoods. As shown in Table 6, location in or near the black ghetto also lowers housing prices, but these effects are not significant at the 5 percent level, at least not with my specification. The second panel of Figure 6 reveals a strikingly similar pattern for neighborhood Hispanic composition. Moving from zero to 49 percent Hispanic lowers the envelope by 37.2 percent, but moving from 49 to 63 percent Hispanic (the maximum composition in the sample) raises the envelope by 2.6 percent. These results support the hypothesis that the Hispanic composition of a neighborhood is a negative amenity for some households and a positive amenity for others.

Table 6 presents results for the geographic controls. Housing prices are significantly higher within two miles of Lake Erie. The coefficients of the two snowbelt variables imply a maximum positive impact on housing prices about 10 miles from Lake Erie in the eastern suburbs, which is exactly where snowfall is the greatest. Housing prices are higher within 10 miles of the airport, but this effect fades to zero by the time 10 miles is reached. Housing prices decline with the dumping of pollutants onto land in the neighborhood and are lower in locations within 20 miles to the northeast or southwest of a major air pollutions site (an effect that declines with distance from the site). Housing prices are boosted by neighborhood amenities and by location near a regional airport, but are lowered by location near a railroad.

Recovering Demand Parameters

The estimated coefficients in Table 5 can be used to calculate ψ for each observation (using equation (19)) and then to estimate equation (13).[40] This step is complicated by the negative values of ψ for some observations for every service or amenity. In three cases (Test Scores, Distance to Hazard, and Non Hispanic), the number of observations with negative values for ψ is small so these observations are simply dropped. In the other cases, however, this procedure does not make sense. Because equation (13) is specified as a multiplicative function, a negative value for ψ implies that one of the variables in the equation must be negative. For example, ψ might depend on the difference between unemployment in a CBG and the average unemployment rate in the sample—a difference that is negative for some CBGs.

This type of model can be estimated with an endogenous switching regression (Lokshin and Sajaia 2004), which has been developed for problems such as estimating wage equations in the public and private sectors.[41] The "switch" in this case is between positive and negative value for ψ. This approach simultaneously estimates three equations: ln(ψ) as a function of household demand traits, Z (at the CBG level), when ψ is positive; ln(-ψ) as a function of Z when ψ is negative; and the sign of the ψ (1=negative; 0=positive) as the realization of a latent variable that depends on Z and other household traits. This model reveals which household traits are associated with a negative value for ψ (that is, with a negative marginal valuation for the associated service or amenity), and then estimates the demand parameters in equation (13) separately for households with positive and negative marginal valuations.

As shown earlier, the income elasticities estimated using equation (13) provide a test of normal sorting, that is, of the hypothesis that $-(\theta/\mu + \gamma) > 0$. Two versions of this test are possible: the standard test based on the net income elasticity from a model with full controls, and a test of equation (29) based on the gross income elasticity from a model with no controls. Results for these tests in Table 7 confirm normal sorting for the two school variables, Safety, and Distance to Hazard. The net income elasticities for locations with negative ψ also confirm normal sorting for High School but contradict it for Safety. Neither of these results remains significant, however, when the impacts of other variables correlated with income are considered.

Finally, this table indicates that among households that prefer a white neighborhood, the marginal willingness to pay for “whiteness” increases with income. Among households that prefer a black neighborhood, however, the gross income elasticity indicates that the marginal willingness to pay for “blackness” declines with income.

The endogenous switching regressions used to derive the results in Table 7 (but not presented separately) also reveal something about the household characteristics significantly associated with downward-sloping bid functions for High School, Safety, and Non Black—and further reinforce the need to consider household heterogeneity. The probability of a negative ψ for High School increases with a CBG’s unemployment rate and its share of students in private school (elementary and high school) and decreases with the share of the CBG’s population that is married, has children, has advanced education, or has a blue collar job. These results are consistent with the view that the value of ψ for High School is most likely to be negative among groups that do not highly value public education. In the case of Safety, the probability of a negative ψ increases with a CBG’s unemployment rate and decreases with the share of a CBGs population that is married, has advanced education, has children, or is elderly. These results are consistent with the view that families and the elderly value safety, but that people who are unemployed have different priorities. Finally, a negative slope for the Non Black is less likely for CBGs in which a relatively large share of the workers have blue collar jobs share or a relatively large share of the students go to private elementary school. These two variables are plausible indicators of prejudice against blacks.

Conclusions

In contrast to Taylor (2008), who writes “Because the hedonic price function is an envelope function, there is no theoretical guidance for its specification” (p. 20), this paper takes the view that the functional form of the housing price function can be derived precisely because it is an envelope. Using constant elasticity demand functions for services/amenities and housing, I derive an individual household's bid function for housing with many service and amenity measures, with varying access to worksites, and with local income and property taxes. Then I derive the housing bid envelope that arises when the underlying demands for services/amenities and housing vary across households. This envelope has a specific form that can be estimated.

This approach has three main advantages. First, it is based on a general characterization of household heterogeneity, which accounts for variation in both observable and unobservable factors in the demand functions for services/amenities and housing. Second, it yields estimates of the price elasticities of demand for services and amenities directly from the hedonic estimation and does not require a Rosen two-step procedure. In other words, once the envelope has been estimated, it is possible to draw the bid functions (and do benefit calculations) for the households receiving any given level of the service or amenity. Third, it provides direct tests of the hypotheses that households sort according to the slopes of their bid functions and that higher-income households win the competition for locations with more desirable amenities.

This approach was implemented using data from Cleveland in 2000. The regressions, which control for a wide range of neighborhood characteristics, indicate a strong impact on housing prices from six different public services or "amenities": school test scores, a measure of high school quality, safety, distance from an environmental hazard, percent non black, and percent non Hispanic. Preliminary tests led me to assume that the price elasticity of demand for safety is -1.0, but price elasticities ranging from -0.13 to -0.62, all precisely estimated, were obtained for the other five variables.

The estimated bid envelopes and underlying bid functions exhibit several different shapes. For Test Scores and Distance to Hazard, moving from the lowest level of the service/amenity to an intermediate level substantially raises housing prices (and underlying bids), but additional movement beyond this point has little additional impact. In the case of the other service/amenities, the impact of amenity increases on housing prices does not fade out at high levels of the service/amenity. Moreover, the results consistently indicate that some households consider these so-called amenities to be disamenities; these households sort into the locations where the "amenity" has the lowest value and lower their housing bids when the "amenity" goes up. This finding demonstrates that household demands are heterogeneous in both observable and unobservable ways and that methods unable to account for this heterogeneity are likely to yield biased results. Additional analysis of the household traits associated with this type of switch in valuation provides support for this interpretation. The households most likely not to value improvements in high school quality, for example, are families without children, with little education, or with children in private school. These results also confirm what surveys have long indicated, namely, that some households prefer largely white neighborhoods, whereas others prefer neighborhoods in which blacks or Hispanics are concentrated.

Finally, the results in this paper provide strong support for the standard model of household bidding and sorting. Results for two "amenities" (high school quality and Hispanic composition) confirm the theorem that households sort into locations according to the slopes of their bid functions. Moreover, the estimated income elasticities confirm the theorem that sorting is "normal," with higher-income households sorting into locations with more desirable services and amenities.

References

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) (January): 20-33.

Bayer, Patrick, Fernando Ferreira, Robert McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy* 115 (4) (August): 558-638.

Bayer, Patrick, and Robert McMillan. 2008. "Distinguishing Racial Preference in the Housing Market: Theory and Evidence." In *Hedonic Methods in Housing Markets: Pricing Environmental Amenities and Segregation*, edited by Andrea Baranzini, José Ramirez, Caroline Schaerer, and Philippe Thalmann (New York: Springer: 2008): 225-244.

Bayer, Patrick, and Christopher Timmins. 2007. "Estimating Equilibrium Models of Sorting Across Locations." *Economic Journal* 117 (518) (March): 353-374

Bayer, Patrick, and Christopher Timmins. 2005. "On the Equilibrium Properties of Locational Sorting Models." *Journal of Urban Economics* 57 (3): 462-77.

Beckman, M. J. 1969. "On the Distribution of Urban Rent and Residential Density." *Journal of Economic Theory* 1 (1) (June): 60-67.

Berry, Steve, Oliver B. Linton, and Ariel Pakes. 2004. "Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems." *Review of Economic Studies* 71 (7): 613-654.

Black, Sandra E. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics* 114 (2) (May): 577-599

Bogart, W.T., and B. Cromwell. 1997. "How Much More is a Good School District Worth?" *National Tax Journal* 70 (2) (June): 215-232.

Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants." *The American Economic Review* 77 (4) (September): 531-553.

Brasington, David M. 2007. "Private Schools and the Willingness to Pay for Public Schooling." *Education Finance and Policy* 2 (2) (Spring): 152-174.

Brasington, David. 2004. "House Prices and the Structure of Local Government: An Application of Spatial Statistics." *Journal of Real Estate Finance and Economics* 29 (2): 211-232.

Brasington. David M. 2002. "The Demand for Local Public Goods: The Case of Public School Quality." *Public Finance Review* 30 (3): 163-187.

Brasington, David M. and Donald R. Haurin. 2006. "Educational Outcomes and House Values: A Test of the Value-Added Approach." *Journal of Regional Science* 46 (2) (May): 245-268.

Brasington, David M., and Diane Hite. 2008. "A Mixed-Index Approach to Identifying Hedonic Price Models." *Regional Science and Urban Economics* 38 (3) (May): 271-284.

Brasington, David M., and Diane Hite. 2005. "Demand for Environmental Quality: A Spatial Hedonic Analysis." *Regional Science and Urban Economics* 35 (1) (January): 57-82.

Brown, James and Harvey Rosen. 1982. "On the Estimation of Structural Hedonic Price Models." *Econometrica* 50: 765-768.

Brueckner, Jan K. 1988. "The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills Model." In *Handbook of Urban and Regional Economics, Vol. II, Urban Economics*, edited by E.S. Mills (Amsterdam: Elsevier): 821-845.

Bui, Linda, and Christopher J. Mayer. 2003. "Regulation and Capitalization of Environmental Amenities: Evidence from the Toxic Release Inventory in Massachusetts." *Review of Economics and Statistics* 85 (3) (August): 693-708.

Charles, Camille Zubrinsky. 2000. "Neighborhood Racial-Composition Preferences: Evidence from a Multiethnic Metropolis." *Social Problems* 47 (3): 379-407.

Chay, Kenneth Y. and Michael Greenstone. 2005. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy* 113 (2): 376-424.

Clapp, John M., Anupam Nanda, and Stephen L. Ross. 2008. "Which School Attributes Matter? The Influence of School District Performance and Demographic Composition on Property Values." *Journal of Urban Economics* 63 (2) (March): 451-456.

Coulson, N. Edward 1991. "Really Useful Tests of the Monocentric Model." *Land Economics* 63 (3) (August): 299-307.

Deng, Yongheng, Stephen L. Ross, and Susan M. Wachter. 2003. "Racial Differences in Homeownership: The Effect of Residential Location." *Regional Science and Urban Economics* 33 (5) (September): 517-556.

Downes, Thomas and Jeffrey E. Zabel. 2002. "The Impact of School Characteristics on House Prices: Chicago 1987-1991." *Journal of Urban Economics* 52 (1) (July): 1-25.

Duncombe, William D., and John Yinger. 1998. "School Finance Reform: Aid Formulas and Equity Objectives." *National Tax Journal*, 51 (2) (June): 239-262.

Ekeland, Ivar, James J. Heckman, and Lars Nesheim. 2004. "Identification and Estimation of Hedonic Models." *Journal of Political Economy* 112 (1, pt. 2): S60-S108.

Ellickson, B. 1971. “Jurisdictional Fragmentation and Residential Choice.” *American Economic Review* 61: 334-339.

Epple, Dennis. 1987. “Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products.” *The Journal of Political Economy* 95 (1) (February): 59-80.

Epple, Dennis, Radu Filimon, and Thomas Romer. 1993. “Existence of Voting and Housing Equilibrium in a System of Communities with Property Taxes.”*Regional Science and Urban Economics* 23 (5) (November): 585-610.

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) (August): 281-308.

Epple, Dennis and G. Platt. 1998. “Equilibrium and Local Redistribution in an Urban Economy When Households Differ by Preferences and Income.” *Journal of Urban Economics* 43 (1) (January): 23-51.

Epple, Dennis, Thomas Romer, and Holger Sieg. 2001. “Interjurisdictional Sorting and Majority Rule: An Empirical Analysis,” *Econometrica* 69 (6) (November): 1437-1465.

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

Follain, James R. and E. Jimenez. 1985. “Estimating the Demand for Housing Characteristics: A Survey and Critique.” *Regional Science and Urban Economics* 15 (1): 77-107.

Gibbons, Steve, and Stephen Machin. 2003. “Valuing English Primary Schools.” *Journal of Urban Economics* 53 (2) (March): 197-219.

Goodspeed, Timothy J. 1989. “A Re-examination of the Use of Ability to Pay Taxes by Local Governments,” *Journal of Public Economics* 38 (3): 319-342.

Glaeser, Edward L., Matthew E. Kahn, and Jordan Rappaport. 2008. “Why Do the Poor Live in Cities? The Role of Public Transportation.” *Journal of Urban Economics* 63 (1) (January): 1–24.

Halvorsen, Robert, and Henry O. Pollakowski. 1981. “Choice of Functional Form for Hedonic Price Equations.” *Journal of Urban Economics* 10 (July): 37-49.

Hartwick, John, U. Schweizer, and P. Varaiya. 1976. “Comparative Statics of a Residential Economy with Several Classes.” *Journal of Economic Theory* 13 (3): 396-413.

Henderson, J. Vernon. 1977. *Economic Theory and the Cities*. New York: Academic Press.

Hoyt, William H., and Stuart S. Rosenthal. 1997. "Household Location and Tiebout: Do Families Sort According to Preferences for Locational Amenities?" *Journal of Urban Economics* 42 (2): 159-178

Ihlanfeldt, Keith R. and B. Scafidi. 2004. "Whites' Neighborhood Racial Preferences and Neighborhood Racial Composition in the United States: Evidence from the Multi-City Study of Urban Inequality." *Housing Studies* (May): 325-359.

Kane, Thomas J., Stephanie K. Riegg, and Douglas O. Staiger. 2006. "School Quality, Neighborhoods, and Housing Prices." *American Law and Economics Review* 8 (2) (Summer): 183-212.

Kiel, Katherine A., and Jeffrey E. Zabel. 1996. "House Price Differentials in U.S. Cities: Household and Neighborhood Racial Effects." *Journal of Housing Economics* 5 (2) (June): 143-165.

Kim, Kyung-Hwan and John F. McDonald. 1987. "Sufficient Conditions for Negative Exponential Densities: A Further Analysis." *Journal of Regional Science* 27 (2) (May): 295-298.

Kim, Chong W, Tim T Phipps, and Luc Anselin. 2003. "Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach." *Journal of Environmental Economics and Management* : 24–39.

Lipman, Barbara J. 2006. "A Heavy Load: The Combined Housing and Transportation Burdens of Working Families." Washington, D.C.: Center for Housing Policy. Available at: http://www.nhc.org.

Lokshin, Michael, and Zurab Sajaia. 2004. "Maximum Likelihood Estimation of Endogenous Switching Regression Models." *The Stata Journal* 4 (3): 282-289.

Mills, Edwin S. 1972. *Urban Economics*. Glenview, IL: Scott Foresman and Company.

Montesano, Aldo. 1972. "A Restatement of Beckman's Model on the Distribution of Urban Rent and Residential Density." *Journal of Economic Theory* 4 (2) (April): 329-354.

Oates, Wallace E. 1969. "The Effects of Property Taxes and Local Public Services on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis." *Journal of Political Economy* 77 (November/December): 957-971.

Ondrich, Jan, Stephen L. Ross and John Yinger. 2003. "Now You See It, Now You Don't: Why Do Real Estate Agents Withhold Available Houses from Black Customers," *The Review of Economics and Statistics*, MIT Press, 85 (4): 854-873.

Polinsky, A.Mitchell and Steven Shavell. 1976. "Amenities and Property Values in a Model of an Urban Area." *Journal of Public Economics* 50 (1-2) (January/February): 119-129.

Pogodzinski, J.M. and D.L. Sjoquist. 1993. "Alternative Tax Regimes in a Local Public Good Model." *Journal of Public Economics* 50 (1): 115-141.

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *The Journal of Political Economy* 82 (1) (January/February): 34-55.

Ross, Stephen L., and Margery Turner, 2005. "Housing Discrimination in Metropolitan America: Explaining Changes between 1989 and 2000." *Social Problems* 52: 152-180.

Ross, Stephen L., Margery Turner, E. Godfrey, and Robin Smith. 2008. "Mortgage lending in Chicago and Los Angeles: A Paired Testing Study of the Pre-application Process." *Journal of Urban Economics* 63: 902-919.

Ross, Stephen, and John Yinger. 1999. "Sorting and Voting: A Review of the Literature on Urban Public Finance." In *Handbook of Urban and Regional Economics, Volume 3, Applied Urban Economics*, edited by P. Cheshire and E. S. Mills (North-Holland): 2001-2060.

Ross, Stephen L., and John Yinger. 1995. "Comparative Static Analysis of Open Urban Models with a Full Labor Market and Suburban Employment." 1995. *Regional Science and Urban Economics* 25(5) (October): 575-605.

Sheppard, S. 1999. "Hedonic Analysis of Housing Markets." In *Handbook of Urban and Regional Economics, Volume 3, Applied Urban Economics*, edited by P. Cheshire and E. S. Mills (North-Holland)

Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf, and Randy Walsh. 2002. "Interjurisdictional Housing Prices in Locational Equilibrium." *Journal of Urban Economics* 52 (1) (July): 131-153.

Small, Kenneth A. Clifford Winston, and Jia Yan. 2005. "Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability." *Econometrica* 73 (4): 1367-1382.

Taylor, Laura O. 2008. "Theoretical Foundations and Empirical Developments in Hedonic Modeling." In *Hedonic Methods in Housing Markets: Pricing Environmental Amenities and Segregation*, edited by Andrea Baranzini, José Ramirez, Caroline Schaerer, and Philippe Thalmann (New York: Springer: 2008): 15-54.

U.S. Department of Transportation, Bureau of Transportation Statistics. 2008. "Table 3-14: Average Cost of Owning and Operating an Automobile." Available at http://www.bts.gov/publications/national_transportation_statistics/html/table_03_14.html

Walsh, Randy. 2007. "Endogenous Open Space Amenities in a Locational Equilibrium." *Journal of Urban Economics* 61 (2) (March): 319-344.

Wieand, Kenneth F. 1987. "An Extension of the Monocentric Urban Spatial Equilibrium Model to a Multicenter Setting: The Case of the Two-center City." *Journal of Urban Economics* 3: 323-343.

Weimer, David L., and Michael J. Wolkoff. 2001. "School Performance and Housing Values: Using Non-Continuous District and Incorporation Boundaries to Identify School Effects." *National Tax Journal* (June): 231-253.

Wheaton, William C. 1993. "Land Capitalization, Tiebout Mobility, and the Role of Zoning Regulations." *Journal of Urban Economics* 34 (2): 102-117.

Wheaton, William C. 1977. "Income and Urban Residence: An Analysis of Consumer Demand for Location." *The American Economic Review* 67 (4) (September): 620-631.

Yinger, John. 1995. "Capitalization and Sorting*:* A Revision." *Public Finance Quarterly*, 23(2) (April): 217-225.

Yinger, John. 1993. "Around the Block: Urban Models with a Street Grid." *Journal of Urban Economics*, 33(3) (May): 305-330.

Yinger, John. 1992. "City and Suburb: Urban Models with More than One Employment Center." *Journal of Urban Economics*, 31(2) (March): 181-205.

Yinger, John. 1982. "Capitalization and the Theory of Local Public Finance." *Journal of Political Economy*, 90(5) (September): 917-943.

Yinger, John. 1979. "Estimating the Relationship between Location and the Price of Housing." 1979. *Journal of Regional Science*, 19 (3) (August): 271-286.

Yinger, John. 1976. "Racial Prejudice and Racial Residential Segregation in an Urban Model." *Journal of Urban Economics* 3(4) (October): 383-396.

Yinger, John, Howard S. Bloom, Axel Boersch-Supan, and Helen F. Ladd. 1988. *Property Taxes and House Values: The Theory and Estimation of Intrajurisdictional Property Tax Capitalization*. Academic Press.

Zabel, Jeffrey A. 2004. "The Demand for Housing Services." *Journal of Housing Economics* 13 (1) (March): 16-35.

Zabel, Jeffrey A. 2008. "Using Hedonic Models to Measure Racial Discrimination and Prejudice in the U.S. Housing Market." In *Hedonic Methods in Housing Markets: Pricing Environmental Amenities and Segregation*, edited by Andrea Baranzini, José Ramirez, Caroline Schaerer, and Philippe Thalmann (New York: Springer: 2008): 177-201.

Endnotes

[1] I prefer the term "ethnic" to "racial," because it emphasizes that distinctions among groups in a society are socially created, even if they draw on superficial physical traits, such as skin color.

[2] This approach, developed in Wheaton (1993), is equivalent to the indirect utility function approach in Epple et al. (1984, 1993).

[3] An informal version of this result is in Yinger (1982), who observes that with sorting the difference in housing prices (per unit of H) between a rich and a poor jurisdiction "does not measure the valuation by either rich or poor households of the difference in service levels between the two jurisdictions" (p. 925).

[4] A person's racial/ethnic attitudes are simultaneously determined with his or her experiences. For a thoughtful demonstration of this simultaneity, see Ihlanfeldt and Scafidi (2004).

[5] This condition holds for any utility function so long as the labor-leisure choice does not alter the value of time as a fraction of the wage (as virtually all urban models assume).

[6] Mills (1972) was the first to use a constant elasticity housing demand function in deriving $P\{u\}$. Kim and McDonald (1987) show that Mills mistakenly uses Y instead of $(Y - tu)$ as the income term. For the sake of consistency, $(Y – tu)$ should be the income concept in (7), too, but this change would lead to the same type of problem that arises if the demand for S_i is a function of S_j. Thus, I assume that households ignore tu in their demand for S (but not for H!).

[7] Many empirical studies include distance or time to work as an explanatory variable in the hedonic equation, without a theoretically derived functional form (Black 1999; Brasington 2002; Downes and Zabel 2002). Exceptions include Yinger (1979) and Coulson (1991).

[8] The Beckman/Montesano approach is equivalent to the derivation of an envelope based on their assumed income distribution, but it cannot accommodate other household traits.

[9] As in the case of commuting distance, changing the income concept in previous derivations to $Y(1\text{-}y)$ would add consistency but also cause problems for the derivation. Because y never exceeds 1 percent, I assume that household demands for S and H can be written without $(1\text{-}y)$.

[10] Similarly, estimation of the σ_T parameters makes it possible to solve equation (27) for ψ_T and to estimate equation (26). This step is not attempted in this paper.

[11] See also Wheaton (1993) and Yinger (1993).

[12] If the terms containing v are moved to the left side, the left side of equations (19), (20), and (28) takes the Box-Cox form. Moreover, if σ_3 is large, the two S terms on the right side of (19) and (28) simplify to a single Box-Cox term.

[13] Several studies also have applied spatial statistics to the estimation of a hedonic equation (Brasington 2004, Brasington and Hite 2005). These techniques usually have little impact on the estimated implicit prices. A few scholars go even farther and argue that a house's price depends on the sales prices of nearby houses (Kim, Phipps, and Anselin 2003).

[14] This approach (1) limits sample size by ignoring houses far from boundaries; (2) controls for omitted variables not obviously correlated with S (because these traits are shared by two locations with different values of S); and (3) fails to recognize that sorting based on S limits the number of shared traits across boundaries (Kane, Riegg, and Staiger 2006. p. 194). Direct controls for neighborhood traits are needed, too (Bayer, Ferriera, and McMillan (2007).

[15] Bajari and Kahn (2005) implicitly consider sorting by allowing marginal willingness to pay for each amenity to vary with household characteristics.

[16] Walsh (2007) uses an approach similar to Epple and Sieg (1999) to estimate the relationship between property values and open space, which is a type of amenity.

[17] Further discussion of their methods can be found in Bayer and Timmins (2005, 2007).

[18] Berry, Linton, and Pakes (2004) show that the method used by Bayer, Ferreira, and McMillan (2007) yields inconsistent estimates unless the number of households increases at a faster rate than the number of houses. Bayer, Ferreira, and McMillan acknowledge this point, but argue that it is not a problem because it is "reasonable" to suppose "that that the number of individuals increases fast enough relative to the number of truly distinct housing types in the market " (p. 634). It is not clear to me why this supposition is reasonable.

[19] This formulation leaves off several terms in the Bayer, Ferreira, and McMillan formulation, their equation (2), including distance to work, their border dummy variables, and unobserved housing or neighborhood traits, all of which can be treated as X's for the purposes of this discussion. Their formulation also includes an error term, which is omitted here for simplicity.

[20] A similar point applies to Bajari and Kahn (2005). Their utility function is linear in the composite good and in the log of amenities. The implied bid function for household type i (with a single X) is $P_i = (Y_i - Ui^0) + \beta_i(\ln\{X\})$. Following the same steps as in the text (with the assumption $\ln\{X_i\} = \sigma_1 + \sigma^2 b_i$) leads to the same envelope with $\ln\{X\}$ replacing X. As a result, the envelope implied by the Bajari and Kahn assumptions contains $\ln\{X\}$ and $[\ln\{X\}]^2$. The squared term is not included in their hedonic equation, which is estimated with local linear regression.

[21] The footnotes to Table 7 indicate that Bayer, Ferreira, and McMillan estimated annualized housing price as a function of the X's, including distance to work and their border fixed effects (their equation (10)). Their discussion on pages 620 and 621 and the title of Table 7 suggest, however, that they used a house-specific fixed effect from their multinomial choice model, called delta, as the dependent variable with price (instrumented) on the right side (their equation (5)). The latter approach yields the MWTP of a household with average values of all the household traits in their data (income, children, race, education), but only if α_{ji} (the coefficient of the

amenity) is divided by α_{hi} (the coefficient of the price variable). Their Table 7 does not present results for the price variable and no mention of this adjustment appears in the article.

[22] Another approach to sorting with household heterogeneity is provided by Hoyt and Rosenthal (1997). They compare hedonic equations that have fixed effects for small neighborhood units with hedonic equations that have the same fixed effects and are differenced over time (to remove household fixed effects). The results of the two approaches are very similar, which is consistent with sorting, that is, with equal household unobservables within neighborhoods.

[23] As explained by Borjas (1987), the standard errors for these second-stage coefficients need to be corrected for heteroskedasticity.

[24] The zip code data comes from the U.S. Census web site. The employment and payroll data are for 2004 (2000 were not available for downloading); the other zip-code data are for 2000.

[25] Many workers do not commute to the worksite center. Ross and Yinger (1995) show that this situation can be analyzed with a housing bid function based on distance or time to the center and a wage gradient that decreases with distance between the center and a worker's job site.

[26] In the single case of a tie, the CBG was assigned to the larger of the two worksites.

[27] I adjust U_i^* until, for each worksite, the two shares differ by less than 1 percentage point.

[28] Bajari and Kahn (2005) and Bayer, Ferriera, and McMillan (2007) use time to work as their commuting variable without considering work location. This approach runs the risk of endogeneity bias whenever unobserved determinants of job location, and hence of distance or time to work, are correlated with housing or neighborhood traits.

[29] Maps are available from the Greater Cleveland Regional Transit Authority: http://www.gcrta.org/maps-schedules.asp.

[30] An appendix explaining these procedures and regression results is available from the author.

[31] Preliminary tests yielded almost identical results with test scores averaged over grades 3, 6, and 9. I also examined a value-added measure made possible by the fact that the data set includes 4th grade scores in 1997-98 and 6th grade scores in 1999-2000, which applied to the same cohort (except, of course, for movement in and out of the district). A preliminary examination of this variable duplicate the finding in Brasington and Haurin (2006; based on the same data) and Downes and Zabel (2002) that this service measure does not affect house values.

[32] I obtained assessment-sales ratios for 2005 to 2008. They appeared to be reasonably accurate: the 2005 ratios averaged 82.2 percent, the ratios were highly correlated across time, and the ratios had little variance (less than 6 percent of the mean). These ratios were not significant in preliminary regressions regardless of how they were averaged over time, so I treat nominal rates as effective rates. The Brasington data contained school tax rates. Other rates and assessment-sales ratios came from http://tax.ohio.gov/channels/research/property_tax_statistics.stm.

[33] The variables indicating distance to public and private schools could be specified as continuous variables and treated as amenities, but preliminary tests found no gain in explanatory power from such an approach compared to using the variables defined in Table 3.

[34] An appendix explaining the data collection procedures is available from the author.

[35] Cleveland snowfall maps are available from the National Weather Service at http://www.erh.noaa.gov/cle/climate/info/snowinfo.html .

[36] If the price elasticity for safety is not assumed to be -1, its estimated value is -0.989 with a very small standard error. However, the SSE is lower when this elasticity is assumed to be -1.

[37] Correcting for the average assessment-sales ratio in 2005 (see note 32) would raise these estimates by about 2 percentage points. In 2000, the national average rate for a fixed-rate 30-year mortgage was 8.06 (http://research.stlouisfed.org/fred2/data/MORTG.txt). Table 3 indicates

an appreciation rate for houses in Cleveland of 0.017 percent per day or $[(1+0.00017)^{365}-1] = 6.4$ percent per year. These figures suggest a real interest rate, r, of 8.06 - 6.4 = 1.66 percent, which I rounded up to 3.0 for the calculations in the text.

[38] These figures use regressions with σ_3=2; figures for other values are virtually the same.

[39] This statement and Figure 6 omit 9 observations with a value of Safety below 0.76. Going from a value of 0.894 to the a value of 0.85 (the lowest value above these outliers) raises housing prices by 1.6 percent; going all the way to the minimum value raises prices by 27.1 percent.

[40] It is also possible to test the condition in (29) by estimating (26) without controls and the value of χ in the text (0.8436). This test supports normal sorting around three worksites, rejects normal sorting for one worksite, and neither supports nor rejects normal sorting for one worksite. Estimating (26) with the controls in the note to Table 7 yields a small, insignificant value for the log of income, which supports the assumption that $\gamma = 1$. In addition, I estimated equation (30) with a Box-Cox specification on the left side (defined as rent using $r = 0.03$ and the tax terms with their estimated coefficients) and right side variables defined by the estimated elasticities presented in the text. The estimated Box-Cox parameter, $(1+v)$ in this case, is -0.124, which indicates $v = -1.124$. This suggests that assuming $v = -1$ may be a reasonable approximation.

[41] An appendix that provides a formal explanation of the endogenous switching regressions along with detailed estimation results is available from the author.

Figure 1: Deriving the Bid-Function Envelope

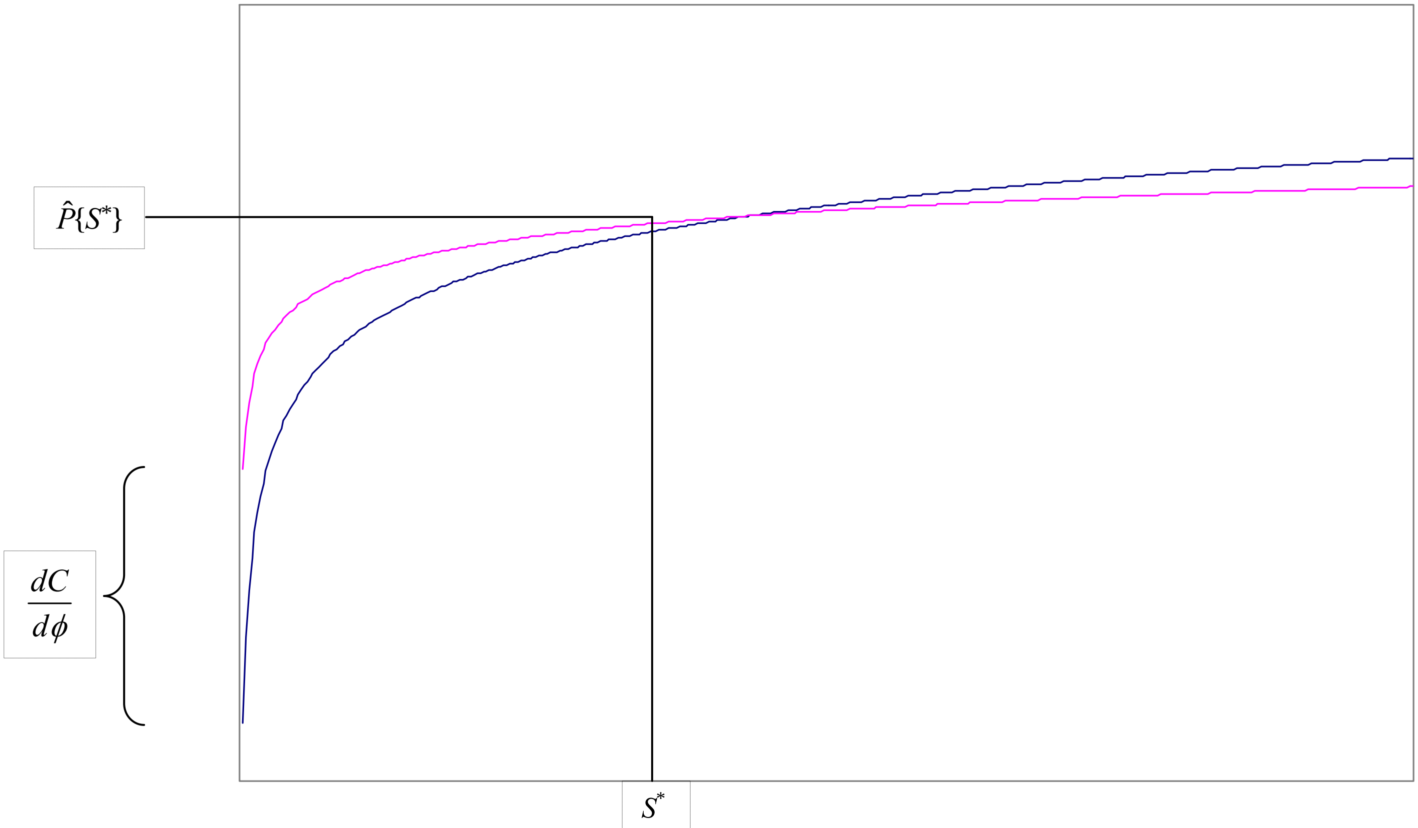

Figure 2. Bid Functions for School Quality

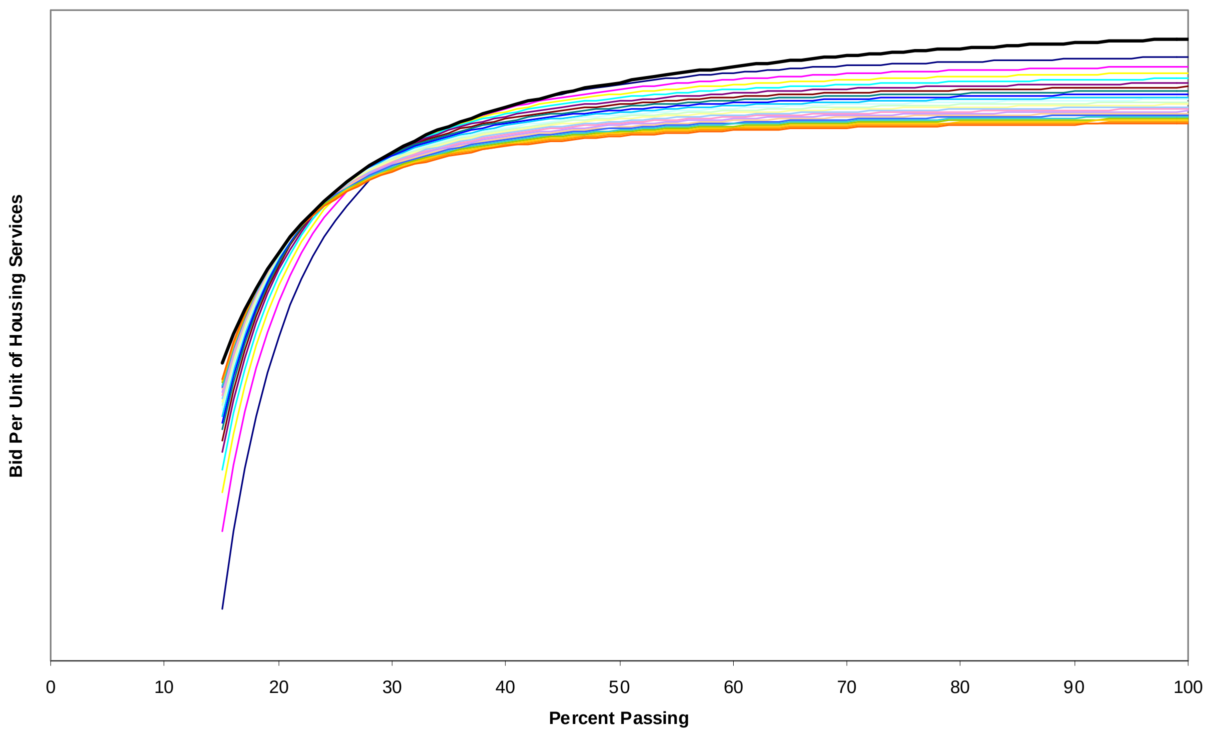

Figure 3. Supply and Demand for School Quality

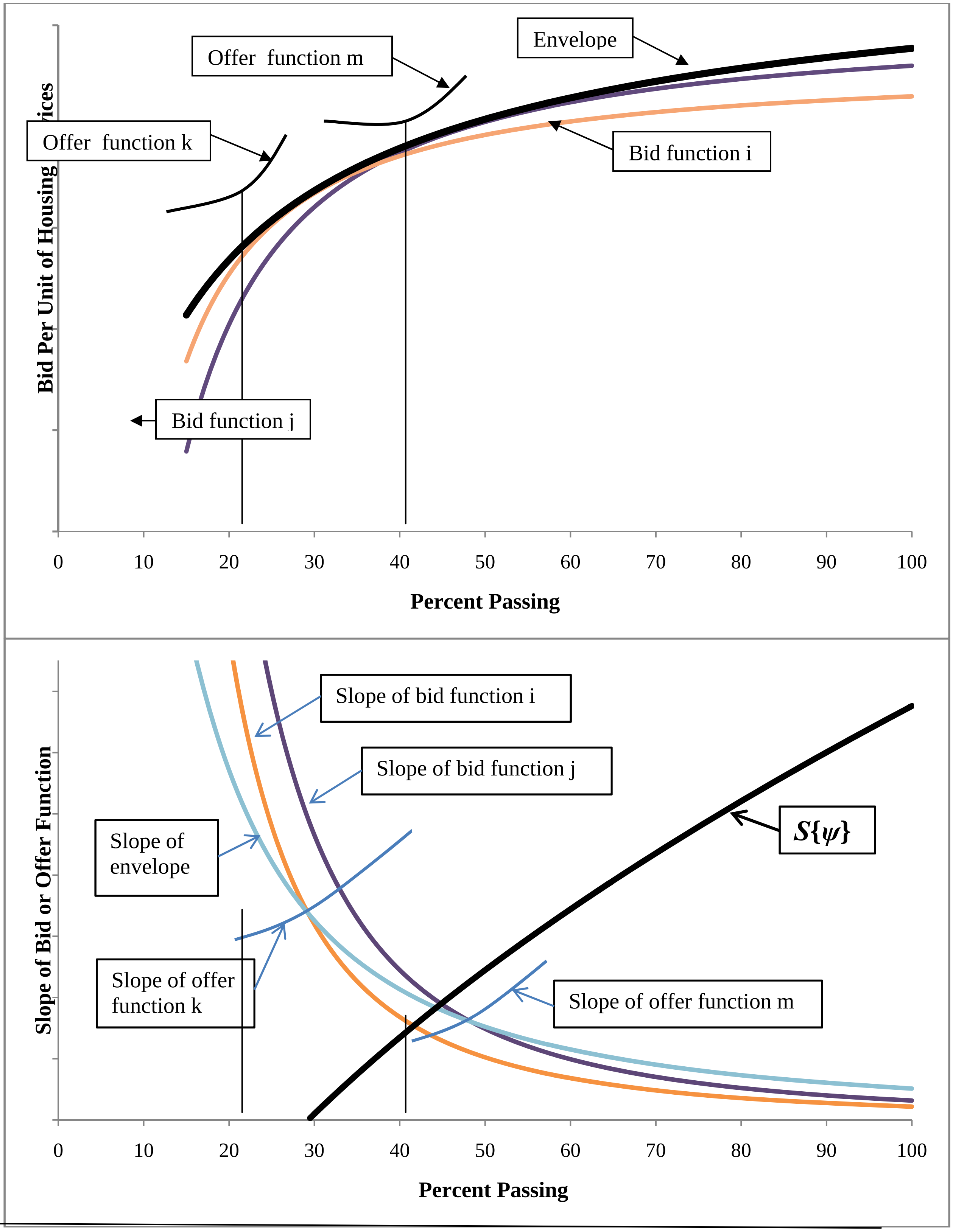

Figure 4. Bid Functions for Distance from Worksite

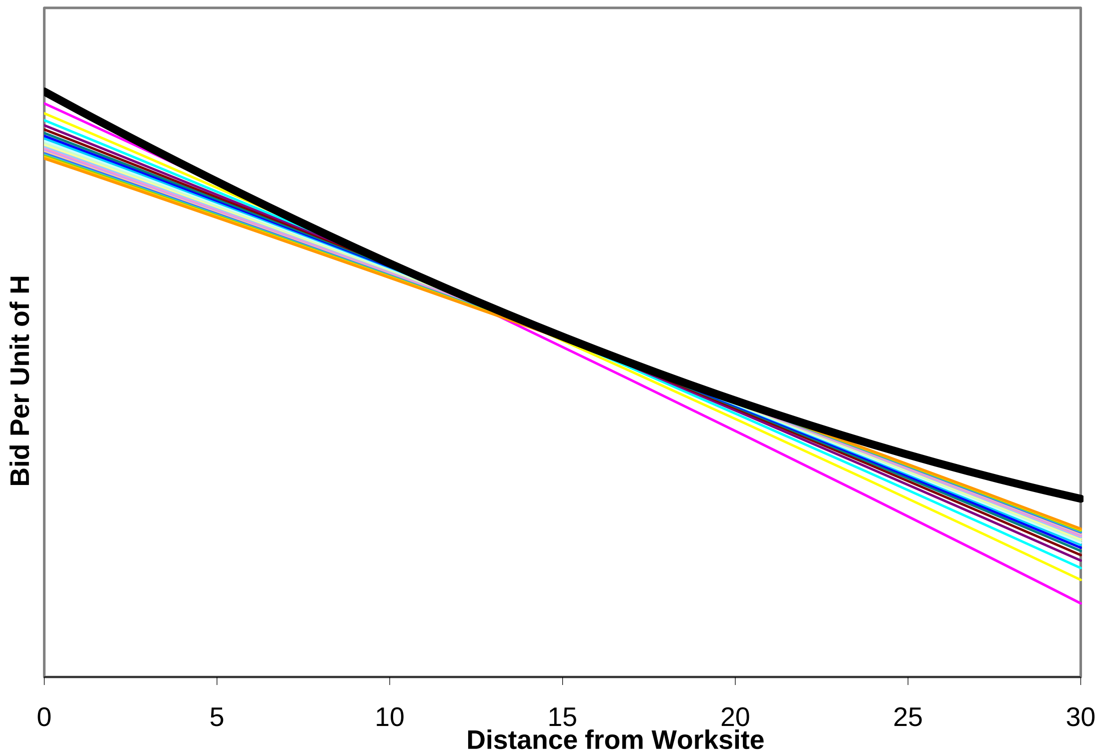

Figure 5. Map of Job Locations and "Worksites" in Cleveland

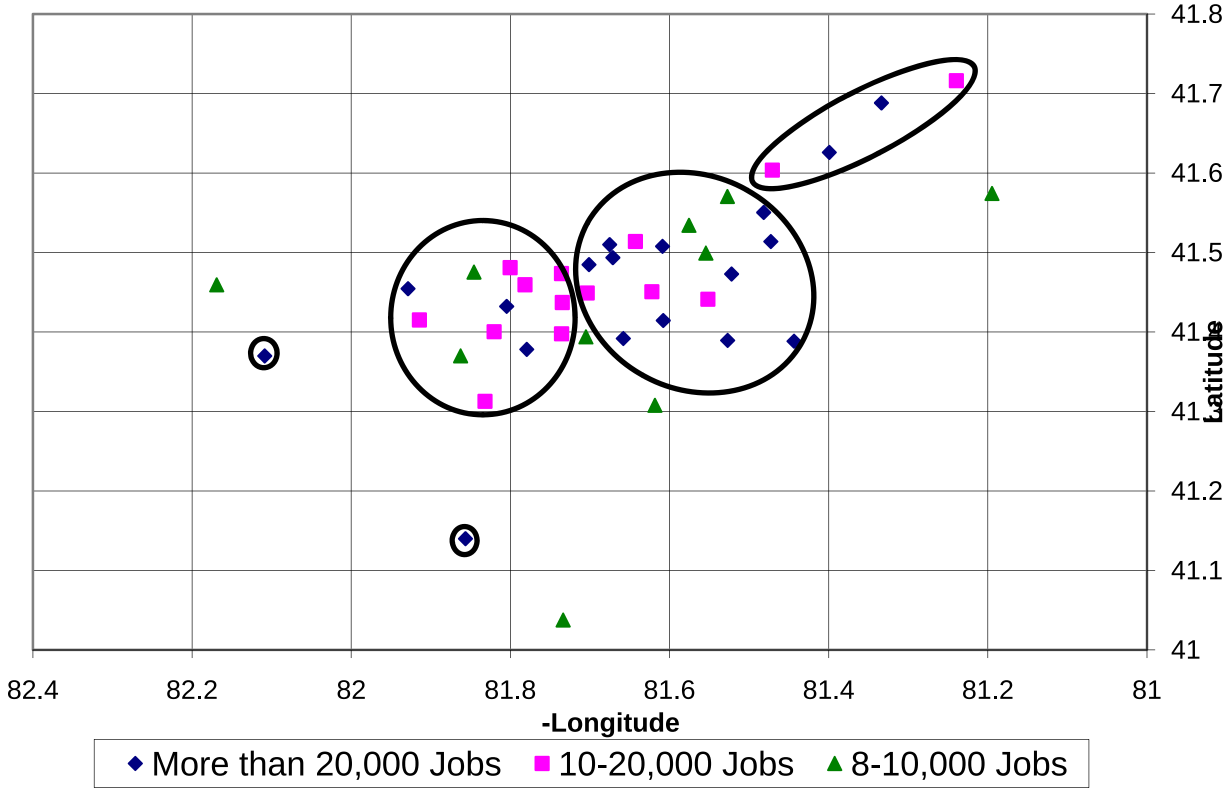

Figure 6. Estimated Envelopes for Public Service and Amenity Variables

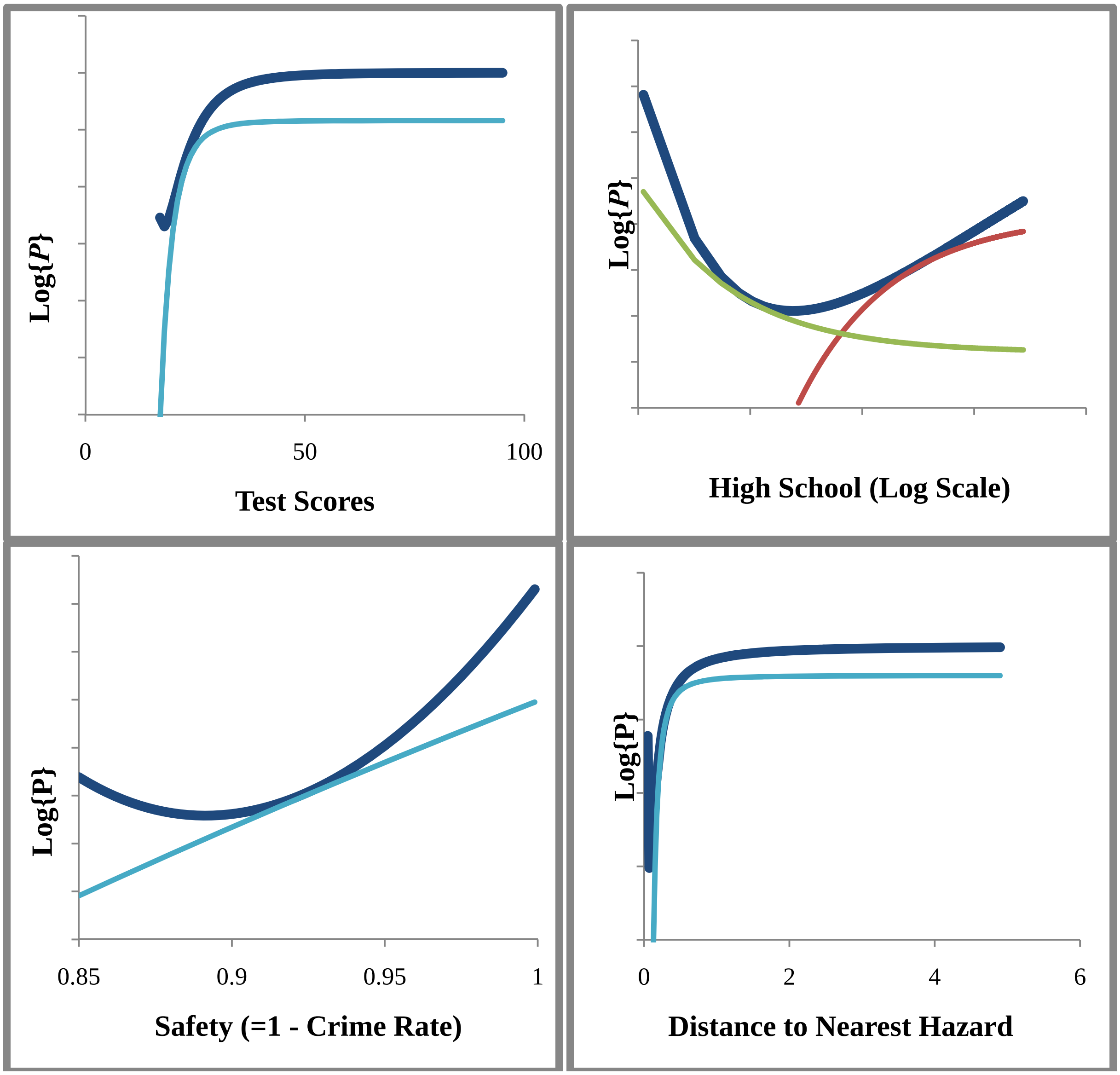

Figure 7. Estimated Envelopes for Ethnicity Variables

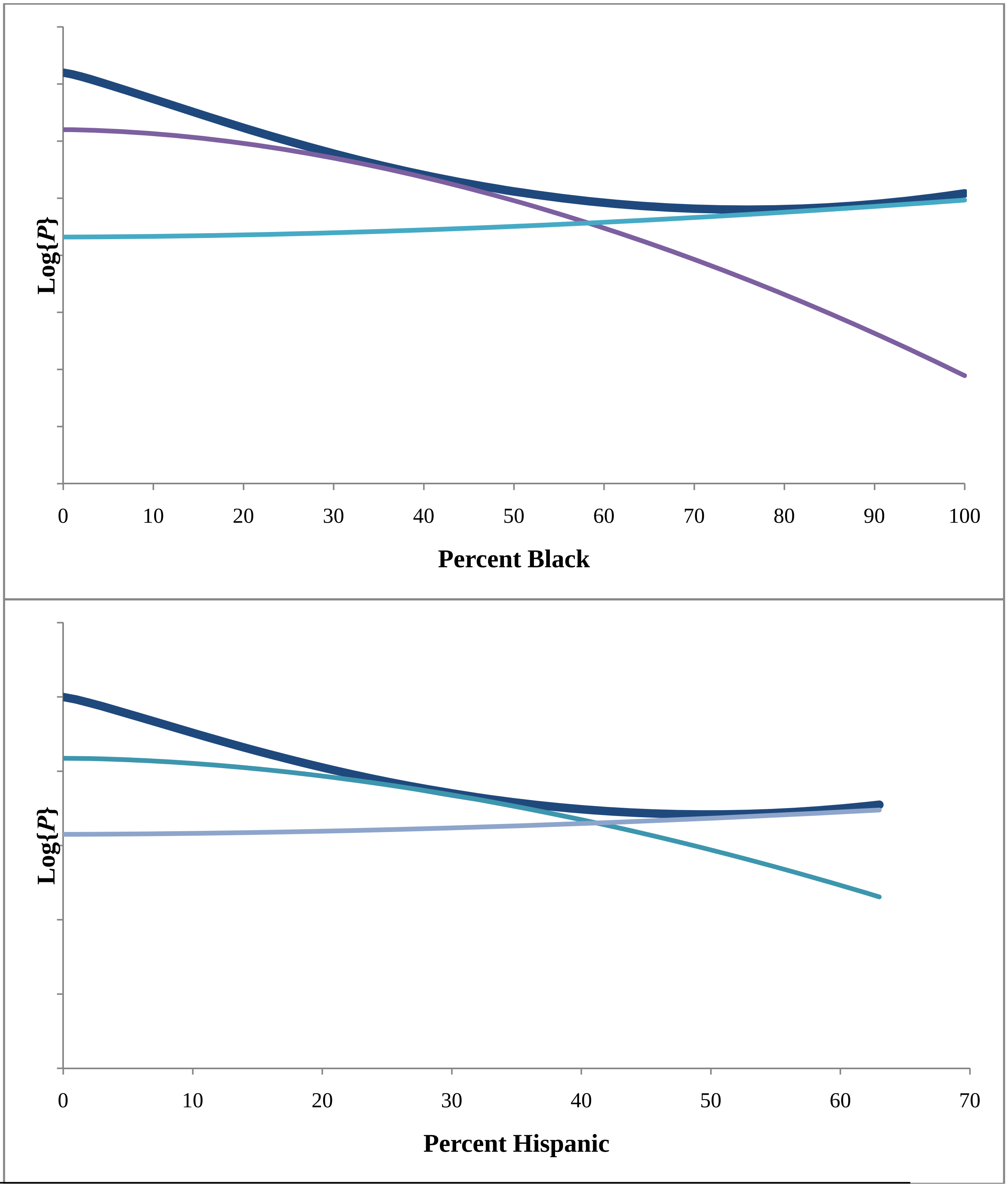

Table 1. Variable Definitions for Basic Hedonic with Neighborhood Fixed Effects

Variable	Definition
One Story	House has one story
Brick	House is made of bricks
Basement	House has a finished basement
Garage	House has a detached garage
Air Cond.	House has central air conditioning
Fireplace	Number of fireplaces
Bedrooms	Number of bedrooms
Full Baths	Number of full bathrooms
Part Baths	Number of partial bathrooms
Age of House	Log of the age of the house
House Area	Log of square feet of living area
Lot Area	Log of lot size
Outbuilding	Number of outbuildings
Porch	House has a porch
Deck	House has a deck
Pool	House has a pool
Date of Sale	Date of house sale (January 1=1, December 31=365)
Commute i[a]	Commuting variable: Log of (Y-tu) for house minus log of (Y-tu) for CBG, worksite i
Public School[a]	Distance to nearest public elementary school (house minus CBG)
Private School	Distance to nearest private school (house minus CBG)
Distance to Hazard	Distance to nearest environmental hazard (house minus CBG)
Distance to Erie[a]	Distance to Lake Erie (if less than 1; house minus CBG)
Distance to Ghetto[a]	Distance to black ghetto (if less than 5; house minus CBG)
Distance to Airport[a]	Distance to Cleveland airport (if less than 10; house minus CBG)

[a] Variable added to the original Brasington data set

Table 2. Results for Basic Hedonic Regression for Cleveland in 2000

Variable	Coefficient	Std. Error	t-stat.	P>|t|
One Story	-0.00673	0.00496	-1.36	0.175
Brick	0.01525	0.00522	2.92	0.004
Basement	0.03154	0.00505	6.24	0.000
Garage	0.13947	0.00673	20.72	0.000
Air Cond.	0.02656	0.00552	4.81	0.000
Fireplaces	0.03162	0.00378	8.37	0.000
Bedrooms	-0.00826	0.00285	-2.90	0.004
Full Baths	0.05997	0.00425	14.10	0.000
Part Baths	0.04175	0.00413	10.10	0.000
Age of House	-0.08288	0.00322	-25.71	0.000
House Area	0.42395	0.00860	49.32	0.000
Lot Area	0.08353	0.00367	22.76	0.000
Outbuildings	0.13706	0.03987	3.44	0.001
Porch	0.03316	0.00727	4.56	0.000
Deck	0.05429	0.00526	10.32	0.000
Pool	0.09191	0.01793	5.13	0.000
Date of Sale	0.00017	0.00002	9.47	0.000
Commute 1	2.98197	0.66112	4.51	0.000
Commute 2	3.07895	0.68387	4.50	0.000
Commute 3	3.08602	0.68817	4.48	0.000
Commute 4	3.20417	0.71337	4.49	0.000
Commute 5	3.13709	0.69455	4.52	0.000
Public School	-0.00181	0.00592	-0.31	0.759
Private School	-0.01575	0.01043	-1.51	0.131
Distance to Hazard	0.02292	0.00588	3.90	0.000
Distance to Erie	-0.01276	0.00485	-2.63	0.008
Distance to Ghetto	-0.17101	0.09026	-1.89	0.058
Distance to Airport	0.08497	0.06294	1.35	0.177
Constant	11.38719	0.01707	667.17	0.000

Notes: Dependent variable = log of transaction amount; 22,857 observations; R^2 = .7899; F(28, 21144) = 455.89 (significant at 0.000 level); 1,685 fixed effects with F(1684, 2144) = 8.45 (significant at 0.000 level); estimated with “areg” command in STATA.

Table 3: Service, "Amenity," Commuting, and Tax Variables

Variable	**Definition**
Test Scores	Average percent passing in school district in 9th grade in 5 state-specified tests (math, reading, writing, science, and citizenship) in 1998-99 and 1999-2000.
High School	Average percent of students in school district with a high pass on the 5 state tests in 12th grade (averaged over 1998-99 and 1999-2000) divided by the drop-out rate (dropouts divided by grade 7-12 enrollment and averaged over 1998-99 and 1999-2000).
Safety	One minus the crime rate (total offenses per 1,000 people in 1997) in CBG (averaged over observations in a CBG, which may be in different police districts).
Hazard	Distance from CBG to nearest environmental hazard
Nonblack	Inverse of percent black in CBG (with 0.01 added to the denominator to account for CBGs with no blacks).
Nonhisp	Inverse of percent Hispanic in CBG (with 0.01 added to the denominator to account for CBGs with no Hispanics).
Commute	Urban model commuting variable for CBG (defined as median family income in CBG minus estimated commuting cost to the worksite assigned to that CBG).
Income Tax Rate	School district income tax rate.
School Tax Rate	School district effective property tax rate.
City Tax Rate[a]	Effective city property tax rate beyond school tax.
Tax Break Rate[a]	Decline in effective total city effective property tax due to exemptions.
Not City[a]	CBG is outside a city (previous two variables = 0).

Notes: In the results tables, a "1" after a variable indicates that the entry is the value of σ_1 for that variable; a "2" indicates that the entry is σ_2; these terms are defined in equation (20) in the text.

[a] Variable added to the original Brasington data set .

Table 4: Geographic Control Variables

Variable	Definition
Lakefront	Center of CBG is within two miles of Lake Erie.
Distance to Lake	Distance in miles from center of CBG to Lake Erie (if less than 2 miles).
Snowbelt 1	Distance in miles from center of CBG to Lake Erie if east of Pepper Pike (and if distance less than 10 miles).
Snowbelt 2	Square of Snowbelt 1
Ghetto	CBG is located in one of the two black ghettos in Cleveland (contiguous areas with more than 80 percent black population).
Near Ghetto	Center of CBG is within 5 miles of the population-weighted center of the nearest ghetto (but not in the ghetto).
Near Airport	Center of CBG is within 10 miles of the center of Cleveland Hopkins Airport.
Airport Distance	Distance in miles from center of CBG to center of Cleveland Airport (if less than 10 miles
Near Public	Center of CBG is one mile or less from nearest public elementary school
Near Private	Center of CBG is one mile or less from nearest private school
Land Release	Pounds of toxic chemicals deposited on land by facilities in the CBG in 2000
Smog d	Center of CBG is within 20 miles of the effluent-weighted geographic centers of one of the three concentrations of air-pollution-emitting facilities in the Cleveland area (in the direction indicated by “d”).
Smog Distance d	Distance in miles from the center of CBG to the effluent-weighted center of the nearest air pollution concentration (if less than 20 miles) (in direction indicated by “d”).
Local Amenities	Number of neighborhood amenities (park, golf course, river, or lake) within 0.25 miles of center of CBG.
Freeway	Center of CBG is within 0.25 miles of a limited-access highway.
Railroad	Center of CBG is within 0.25 miles of a railroad.
Shopping	Center of CBG is within 1 mile of a shopping center or mall.
Hospital	Center of CBG is within 1 mile of a hospital
Small Airport	Center of CBG is within 1 mile of a small airport
Big Park	Center of CBG is within 1 mile of a state, regional, or county park (or the Cleveland zoo).

Notes: Except for “Land Release” and “Near Private,” all of these variables were added to the original Brasington data set by the author, although they were all derived using the CBG latitudes and longitudes in that data set and, in the case of the ghetto and smog variables, on other information in the data set, too.

Table 5: Results for Service, "Amenity," Commuting, and Tax Variables

	Standard Model[a]	All μ's equal -1[b]	Estimated μ's[c] σ_3 =1	σ_3 =2	σ_3 =9
Test Scores 1	0.00482***	0.25180*	18.03901***	4.22993***	1.37705***
	(10.92)	(1.98)	(89.30)	(94.88)	(337.88)
Test Scores 2		-0.03106	8.33E-11	6.54E-10	0.00000
		(-0.81)	(0.16)	(0.20)	(0.22)
Test Scores mu		-1.00000	-0.11435***	-0.13369***	-0.14411***
			(-4.39)	(-4.95)	(-5.00)
High School 1	0.00119***	-0.04363**	2.65250***	1.70458***	1.13036***
	(4.41)	(-2.34)	(5.08)	(9.94)	(48.23)
High School 2		0.03624***	15.98304***	6.83362***	1.26833***
		(4.58)	(2.47)	(2.60)	(2.73)
High School mu		-1.00000	-0.51434***	-0.62335***	-0.72511***
			(-13.74)	(-11.50)	(-10.18)
Safety 1	0.21327	-30.72455***	0.89291***	0.94415***	0.98716***
	(0.84)	(-3.8)	(61.35)	(121.98)	(557.78)
Safety 2		130.18362***	0.06208***	0.03310***	0.00764***
		(3.79)	(3.38)	(3.38)	(159.02)
Safety mu		-1.00000	-1.00000	-1.00000	-1.00000
Distance to Hazard 1	0.02819***	0.09062***	0.08076***	0.28467***	0.75646***
	(3.96)	(3.59)	(3.27)	(6.24)	(27.60)
Distance to Hazard 2		-0.14275***	70.39408	55.86535	19.15127
		(-2.58)	(1.63)	(1.62)	(1.63)
Distance to Hazard mu		-1.00000	-0.28952***	-0.32046***	-0.34504***
			(-6.04)	(-5.23)	(-4.79)
Non Black 1	0.00020**	0.04947***	0.01321***	0.11482***	0.61808***
	(2.06)	(8.5)	(8.32)	(15.32)	(66.79)
Non Black 2		-0.02742***	241.15510*	242.79090*	110.27900*
		(-7.78)	(1.99)	(1.88)	(1.86)
Non Black mu		-1.00000	-0.31757***	-0.35798***	-0.39012***
			(-17.13)	(-14.03)	(-12.45)
Non Hispanic 1	0.00024***	0.04172***	0.02050***	0.14175***	0.64715***
	(2.77)	(6.05)	(8.26)	(13.69)	(57.11)
Non Hispanic 2		-0.02115***	115.91550***	106.18720***	44.52483***
		(-5.15)	(2.75)	(2.62)	(2.61)
Non Hispanic mu		-1.00000	-0.31543***	-0.35824***	-0.39196***
			(-18.95)	(-15.13)	(-13.26)
Commute 1	20.32068***	17.34504***	0.18450**	0.18779**	0.18872**
	(3.15)	(2.69)	(2.11)	(2.19)	(2.20)
Commute 2	-21.18214***	-17.91119**	0.20542***	0.20411***	0.20395888***
	(-2.99)	(-2.52)	(2.95)	(2.97)	(2.96)

Income Tax Rate	1.53928	-4.90929	-6.41500*	-6.22136*	-6.15772*
	(0.45)	(-1.4)	(-1.77)	(-1.72)	(-1.73)
School Tax Rate	1.42261	-0.86179	1.11980	1.06216	1.04468
	(1.2)	(-0.68)	(0.90)	(0.86)	(0.87)
City Tax Rate	-2.13914*	-3.10726***	3.78116***	3.71210***	3.69267***
	(-1.86)	(-2.57)	(2.75)	(2.70)	(2.73)
Tax Break Rate	3.06068*	5.76055***	-6.91241***	-6.84451***	-6.82630***
	(1.81)	(3.23)	(-3.35)	(-3.32)	(-3.35)
Not City	-0.00850	0.01034	0.03243	0.03351	0.03355
	(-0.32)	(0.38)	(1.16)	(1.20)	(1.21)
SSE	43.9972	39.2417412	37.31037	37.29866	37.29533
R-squared	0.6222	0.663	0.6796	0.6685	0.6687
Root MSE	0.16394	0.15511	0.1514332	0.1514094	0.1513561

Notes: The dependent variable is the CBG fixed effect (in log form); the number of observations is 1685; the regressions in this table also include the variables in Table 6 plus worksite and county fixed effects.

[a]The results in this column are based on linear version of the service and "amenity" variables.

[b]This regression sets σ_3 = 2.

[c] These regressions set μ = -1 for the safety variable

Table 6: Results for Geographic Control Variables

	Standard Model[a]	All μ's equal -1[b]	Estimated μ's[c] $\sigma_3=1$	$\sigma_3=2$	$\sigma_3=9$
Lakefront	0.112646***	0.08041***	0.07160***	0.07181***	0.07182***
	(4.01)	(3.01)	(2.630)	(2.64)	(2.64)
Distance to Lake	-0.07162***	-0.04217**	-0.03210	-0.03260	-0.03280
	(-3.39)	(-2.09)	(-1.580)	(-1.61)	(-1.62)
Snowbelt 1	0.038168***	0.04019***	0.03903***	0.03899***	0.03901***
	(4.28)	(4.48)	(4.260)	(4.27)	(4.37)
Snowbelt 2	-0.00168***	-0.00191***	-0.00190***	-0.00190***	-0.00191***
	(-3.25)	(-3.63)	(-3.540)	(-3.55)	(-3.67)
Ghetto	-0.11216***	-0.08443***	-0.06626*	-0.06511*	-0.06463*
	(-4.81)	(-3.49)	(-1.880)	(-1.86)	(-1.85)
Near Ghetto	-0.05815***	-0.04215*	-0.04444*	-0.04352*	-0.04316*
	(-2.74)	(-1.95)	(-1.930)	(-1.89)	(-1.88)
Near Airport	0.176843***	0.16758***	0.15121***	0.15155***	0.15170***
	(4.81)	(4.77)	(4.250)	(4.25)	(4.33)
Airport Distance	-0.02509***	-0.02448***	-0.01945***	-0.01962***	-0.01968***
	(-5.9)	(-6.35)	(-5.070)	(-5.10)	(-5.14)
Near Public	-0.00014	-0.00171	-0.00364	-0.00364	-0.00361
	(-0.01)	(-0.16)	(-0.330)	(-0.33)	(-0.33)
Near Private	0.028782*	0.02214*	0.01497	0.01458	0.01440
	(2.12)	(1.75)	(1.180)	(1.15)	(1.14)
Land Release	-5.28E-06***	0.00000**	0.00000***	0.00000***	0.00000***
	(-5.65)	(-2.48)	(-2.860)	(-2.75)	(-2.71)
Smog NE	-0.23695***	-0.23274***	-0.33621***	-0.33474***	-0.33327***
	(-2.88)	(-2.81)	(-3.980)	(-3.97)	(-3.92)
Smog SE	-0.14291*	-0.09371	-0.18592***	-0.18607*	-0.18510**
	(-1.8)	(-1.22)	(-2.220)	(-2.23)	(-2.20)
Smog SW	-0.18347**	-0.16818**	-0.22779***	-0.22654***	-0.22524***
	(-2.5)	(-2.15)	(-2.790)	(-2.77)	(-2.75)
Smog NW	0.04848	0.14579	0.01803	0.02100	0.02291
	(0.61)	(1.39)	(0.130)	(0.15)	(0.16)
Smog Direction NE	0.01167***	0.01134**	0.01704***	0.01696***	0.01689***
	(2.63)	(2.53)	(3.720)	(3.70)	(3.65)
Smog Direction SE	0.004248	0.00341	0.00976*	0.00984*	0.00985*
	(0.78)	(0.67)	(1.730)	(1.75)	(1.76)
Smog Direction SW	0.007591*	0.00784*	0.01078**	0.01079**	0.01074**
	(1.75)	(1.66)	(2.200)	(2.20)	(2.18)
Smog Direction NW	-0.09881***	-0.13551***	-0.08712	-0.08775	-0.08801
	(-3.26)	(-2.93)	(-1.390)	(-1.41)	(-1.41)
Local Amenities	0.018543**	0.01806**	0.01644**	0.01655**	0.01661**
	(2.19)	(2.25)	(2.040)	(2.05)	(2.06)
Freeway	0.004241	0.00656	0.00782	0.00763	0.00759
	(0.33)	(0.55)	(0.660)	(0.64)	(0.64)

Railroad	-0.04067***	-0.02422**	-0.02098**	-0.02085**	-0.02078**
	(-3.72)	(-2.26)	(-1.990)	(-1.97)	(-1.96)
Shopping	0.008763	0.00613	0.00309	0.00317	0.00322
	(0.89)	(0.66)	(0.340)	(0.34)	(0.35)
Hospital	-0.01973*	-0.02002*	-0.01357	-0.01384	-0.01393
	(-1.81)	(-1.91)	(-1.290)	(-1.32)	(-1.33)
Small Airport	0.040574*	0.04785**	0.05839**	0.05835**	0.05832**
	(1.66)	(2.13)	(2.470)	(2.48)	(2.48)
Big Park	0.038692***	0.01566	0.01623	0.01624	0.01627
	(3.38)	(1.46)	(1.570)	(1.57)	(1.57)

Notes: The dependent variable is the CBG fixed effect (in log form); the number of observations is 1685; the regressions in this table also include the variables in Table 5 plus worksite and county fixed effects.

[a]The results in this column are based on linear version of the service and "amenity" variables.

[b]This regression sets σ_3 = 2.

[c] These regressions set μ = -1 for the safety variable.

Table 7. Tests of the Normal Sorting Hypothesis

	Test Scores	High School	Safety	Hazard	Non Black	Non Hispanic
RESULTS WITH FULL CONTROLS (NET INCOME ELASTICITY)						
Log of Family Income	0.84539***	1.034509***	0.21305***	0.1765**	1.224279***	0.170024
	(7.81)	(3.71)	(2.90)	(2.27)	(5.55)	(0.391)
Log of Family Income if Slope of Bid Function Is Negative	-0.18157***	-0.1262513**	0.13928**		-0.13799	
	(-12.01)	(-2.28)	(1.98)		(-0.59)	
RESULTS WITH NO CONTROLS (GROSS INCOME ELASTICITY)						
Log of Family Income	1.91532***	2.230293***	0.46413***	0.769185***	1.675584***	0.868852***
	(35.23)	(4.22)	(9.74)	(23.46)	(13.67)	(8.41)
Log of Family Income if Slope of Bid Function Is Negative		-0.0807648	0.17831		-0.39031**	
		(-0.22)	(1.24)		(-2.42)	
Number of Observations	1655	1685	1685	1678	1685	1671

Notes: The dependent variable is the slope of the bid function (ψ); a few observations with negative values for ψ are dropped from the regressions in columns 1, 4, and 6; the regressions in columns 2, 3, and 5 are estimated as endogenous switching regressions; the results in the first panel use the following variables as controls: the percentage of households in the CBG that have children are married, just speak English at home, or are owner-occupants; the percentage of the population in the CBG that is married, is foreign-born, is Asian or Pacific Islander, has a high school degree only, has some college, has a college degree, or has a graduate degree; the percentage of the labor force that is unemployed; the percentage of workers with a blue collar job; the percentage of households in the county who are Catholic or who belong to a non-catholic church; the percentage of elementary students in private school; the percentage of high school students in private school; and the percentage of households in the tract who moved within the last year.