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

John Yinger

Center for Policy Research

Syracuse University

May 2009

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 Jim Follain, Eric Isenberg, Phuong Nguyen, Chris Rohlfs, Steve Ross, and participants in a seminar at Syracuse University.

Author's address: Center for Policy Research, Eggers Hall, Syracuse University, Syracuse, NY 13244.

Author's e-mail: jyinger@maxwell.syr.edu

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 the Cleveland area in 2000 yields precise estimates of price elasticities for school quality, the crime rate, 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 lowering bids as some "amenities" increase.

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 service and amenity demand elasticities.

The foundation of this paper is the theory of household bids for housing in locations with different characteristics. The approach developed here brings together bids for housing based on public services, neighborhood amenities, commuting costs, and local taxes. Two key steps define this approach. 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 a specific form for the bid functions that can accommodate many different public services and amenities.

Second, I introduce household heterogeneity into the public service, amenity, and housing demand functions. Using a 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 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, a feature that proves to be important empirically.

The second part of the paper estimates this house value regression for all the house sales in the Cleveland area in 2000. I find strong impacts on house values from several school district characteristics, the crime rate, and distance from environmental hazards, along with 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 section provides a brief introduction to the relevant literature; a full discussion is postponed until a later section to facilitate comparisons with my approach.

The foundation of this paper is 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 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 literature predicts that property taxes and public service levels are 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 Bayer et al. (2007), Black (1999), Brasington (2002, 2004, 2007); Brasington and Haurin (2006); Bogart and Cromwell (1997); Clapp, Nanda, and Ross (2008); Downes and Zabel (2002); Figlio and Lucas (2004); 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 literature concerns hedonic regressions, 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 to explore 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 regression. The second step is to find implicit prices of product attributes by differentiating the hedonic regression and then to estimate the demand for each attribute as a function of its implicit price.

Brown and Rosen (1982) and Epple (1987) explain 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. Ekeland et al. also identify a "problem with the current literature," namely, that the "full economic content of the hedonic model is not being exploited" (p. S74). This paper takes this lesson to heart by exploiting the economic content in bidding and sorting models to derive a hedonic price function with which the service/amenity price elasticities can be directly estimated, without the need for a second step.

Sorting is brought into a hedonic regression by Epple and Sieg (1999) and Epple, Romer, and Seig (2001). These studies estimate a general equilibrium model in which households sort into communities and then select the level of local public services. Bayer et al. (2007) address sorting using a discrete choice model in which households are allocated across houses according their valuation of housing and neighborhood attributes

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 (1987), 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 bidding framework based on the standard model with the addition of 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.

Housing Bids and Locational Equilibrium with Public Services

The 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 households are mobile, so a key equilibrium condition is that all households in an income-taste class achieve the same utility. 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. 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.

More formally, households maximize utility 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 (and $\tau^* = \tau/r$). Thus,

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

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 (1) for *P*, this approach leads to the following maximization problem:

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

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. The envelope theorem can be used to determine the impact of *S* and τ on bids. Specifically,

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

and

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

where subscripts indicate partial derivatives. The numerator of (3) is the marginal rate of substitution between *S* and *Z*, 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 a 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, which has been widely used in demand functions estimated with community-level data (Duncombe and Yinger 1998), 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*. The associated inverse demand function for *S* is:

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

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

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

where K_H is a constant, M is a vector of preference and other variables that influence housing demand, and $\hat{P}$ is defined by (5). Combining (3), (6), (8), and (9) yields

$$\hat{P}_S \hat{P}^{\nu} = \frac{S^{1/\mu}}{\left(K_S N^{\delta}\right)^{1/\mu} K_H M^{\rho} Y^{(\theta/\mu)+\gamma}} = \psi S^{1/\mu} \ , \tag{10}$$

where

$$\psi = \left(\left(K_S N^{\delta}\right)^{1/\mu} K_H M^{\rho} Y^{(\theta/\mu)+\gamma}\right)^{-1} . \tag{11}$$

Equation (10) is an exact differential equation, which can be solved by integrating both sides:

$$\hat{P}\{S\} = \left(\left(1+\nu\right)\left(C + \left(\frac{\psi\mu}{1+\mu}\right) S^{(1+\mu)/\mu}\right)\right)^{1/(1+\nu)} , \tag{12}$$

where C is a constant of integration. Note that if $\nu = 1$, this equation simplifies to

$$\ln\{\hat{P}\{S\}\} = C + \left(\frac{\psi\mu}{1+\mu}\right) S^{(1+\mu)/\mu} . \tag{13}$$

Sorting

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 et al. (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 et al. 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*. These models all rely on 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 (10). A household's relative slope is determined by all the variables in (10) that are unique to that household, that is, everything except for *S* and $\hat{P}$. These terms are collected in ψ as defined by (11). In other words, ψ can be interpreted as in index of the steepness of a household's bid function—and hence an 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, the bid function with the steeper slope must have a smaller intercept. The derivation of an envelope therefore involves finding a formula for *C* 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 (12) yields

$$\left.\frac{dC}{d\psi}\right|_{S=S\{\psi\}} = \frac{-\mu}{1+\mu} S\{\psi\}^{\frac{1+\mu}{\mu}} . \tag{14}$$

To solve this differential equation, that is, to find the envelope, we need to specify a form for $S\{\psi\}$.[3] 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. A reasonably general form that leads to a differential equation with an analytical solution is

$$S\{\psi\} = (\sigma_1 + \sigma_2\psi)^{\sigma_3} \ . \tag{15}$$

Substituting (17) into (16) and integrating yields:

$$C = C_0 - \left(\frac{\mu}{1+\mu}\right)\left(\frac{\mu}{\sigma_3(1+\mu)+\mu}\right)\left(\frac{(\sigma_1+\sigma_2\psi)^{\frac{\sigma_3(1+\mu)+\mu}{\mu}}}{\sigma_2}\right), \tag{16}$$

where C_0 is a constant of integration. Solving (15) for ψ and substituting the result and (16) into (12) leads to the bid envelope, identified with superscript E:

$$\hat{P}^E\{S\} = \left[(\nu+1)\left(C_0 - \left(\frac{\sigma_1}{\sigma_2}\right)\left(\frac{\mu}{1+\mu}\right)S^{\left(\frac{1+\mu}{\mu}\right)} + \left(\frac{1}{\sigma_2}\right)\left(\frac{1}{\frac{1+\mu}{\mu}+\frac{1}{\sigma_3}}\right)S^{\left(\frac{1+\mu}{\mu}+\frac{1}{\sigma_3}\right)}\right)\right]^{1/(\nu+1)} \ . \tag{17}$$

Equation (17) makes it clear that the impact of S on bids depends not only on a household's willingness to pay for S, as indicated by μ, but also on sorting, as measured by the σ parameters. Thus, 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.[4]

Further explanation of this approach is provided by Figure 2, which is an extension of figures in Follain and Jimenez (1985) and Taylor (2008). The bid-function envelope, illustrative bid functions for household types, and 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 the bottom panel. For ease of interpretation, 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 (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

Under fairly general assumptions, this approach can be applied to many public services and to amenities, such as clean air. Let S_i be the ith public service or amenity (henceforth called an "amenity" for conciseness) with its own constant-elasticity demand function. If the demand for S_i does not depend on any S_j, $j \neq i$, then the above expressions for S_j can be treated as part of the constant of integration, C, in (17) for S_i, and the envelope is

$$\hat{P}^E\{S_1,\ldots,S_n\} = \left((\nu+1)\left(C_0 + \sum_i \left(\left(\frac{1}{\sigma_{2i}\eta_i}\right)(S_i)^{\eta_i} - \left(\frac{\sigma_{1i}}{\sigma_{2i}\zeta_i}\right)(S_i)^{\zeta_i} \right)\right)\right)^{1/(\nu+1)}, \tag{18}$$

where

$$\eta_i = \left(\frac{1+\mu_i}{\mu_i} + \frac{1}{\sigma_{3i}}\right) \quad \text{and} \quad \zeta_i = \left(\frac{1+\mu_i}{\mu_i}\right). \tag{19}$$

In standard demand theory, the first-order conditions for utility maximization 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 above model, 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, and the amount that a household bids to live in a community with a certain value of S_i might depend on how much S_j the community provides. If so, the derivative of the household-type-specific bid function with respect to S_j has two terms, (10) and a new term representing the impact of a change in S_j on the household's willingness to pay for S_i. Differentiating (12) no longer leads to the same answer as integrating (10). 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 (7) for S_i, and therefore cause no complications for the derivation.[5]

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 (Charles 2000), so prejudice alone is unlikely to generate a monotonic relationship between a neighborhood's ethnic composition and housing price.[6]

Second, historically disadvantaged minority groups tend to cluster in neighborhoods with poor amenities, even 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 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 in certain ethnic groups might face discrimination in obtaining housing or mortgages. Although the extent of discrimination has declined over time, recent audit studies find evidence of continuing discrimination against blacks and Hispanics in these markets (Ross and Turner 2005; Ross et al. 2008). Discrimination could push blacks and Hispanics into neighborhoods with poor amenities, thereby reinforcing the correlation between unobserved neighborhood traits and ethnicity. One study (Ondrich, Ross, and Yinger 2003) found that after controlling for housing price, real estate brokers steered blacks, but not whites, toward neighborhoods with lower average house values or higher average house ages than the one originally requested. Discrimination also might restrict the ability of minority households to move out of largely minority neighborhoods and thereby artificially boost housing prices there.[7]

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 (17), a household's marginal willingness to pay for an "amenity," ψ, is negative if

$$(\sigma_1)^{\sigma_3} \equiv \sigma^* > S \, . \tag{20}$$

Any bid function that is tangent to the envelope at a value of S below σ^* has a negative slope.

The regressions presented below include two ethnicity-based “amenities”: the inverse of the black or Hispanic concentration in a neighborhood. People who prefer a white, non-Hispanic neighborhood place a positive value on these “amenities.” If some households (of any ethnicity) value an increase in the black (or Hispanic) share, then the estimate of σ^* will be greater than the minimum value of this “amenity” variable in the sample. Unlike previous work on this topic (including my own), therefore, this approach does not confuse the estimated envelope with the willingness to pay of a given group. Instead, heterogeneity in household preferences is revealed by the estimation. The regressions also identify neighborhoods in or near a black ghetto (contiguous neighborhoods at least 80 percent black). These variables control for unmeasured neighborhood quality variables in the ghetto, discriminatory restrictions on the ability of blacks to move outside the ghetto, and expectations about potential racial transition near the ghetto.

Commuting Costs

A standard urban model specifies housing bids as a function of distance to a worksite (Brueckner 1987). 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.

An alternative approach is to derive 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. This envelope is derived in Yinger (2009). The empirical work presented below includes two commuting variables that are analogous to the variables in

(17), under the assumptions that the $\gamma = 1$ and that the relationship between bid-function slopes and commuting distance, which is analogous to (15), can be approximated with a linear function.

The Final Estimating Equation

Finally, following (1) and (5) 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:

$$V = \frac{P^E H\{X\}}{r} = \frac{\hat{P}^E\{S_1, S_2, \ldots, S_n, u\}H\{X\}}{(r + \beta\tau)}, \tag{21}$$

where $\hat{P}^E$ reflects (18) and the comparable terms for commuting costs. I specify $H\{X\}$ as multiplicative; Sieg et al. (2002) show that this form is consistent with the assumption used here 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 (18) (through (21)) yields μ each amenity. Moreover, estimation of the σs makes it possible to solve (15) for ψ and hence to estimate (11). Because S does not appear on the right side of (11), this estimation avoids the endogeneity problem that arises with the Rosen two-step method. One disadvantage of this approach is that it cannot generally separate the impacts on S and H of a variable that appears in both demand functions.

"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 a amenity requires that $\theta/\mu + \gamma < 0$, so the coefficient of the income term in (11) provides a test of this condition. See also Wheaton (1993) and Yinger (1993).

Comparison with Previous Approaches

In Rosen (1974) the hedonic price function is "a joint envelope of a family of value functions and another family of offer functions" (p. 44). Some scholars derive joint envelopes that are consistent with equilibrium when the value (or bid) functions and offer functions take certain forms (Epple 1987; Ekeland et al. 2004). As Ekeland et al. point out it, however, this "approach is computationally complicated and does not transparently deliver identification of the deep structural parameters" (p. S71). A fundamental problem for any hedonic price study, therefore, is how to specify the regression.

Many scholars estimate the hedonic price function using general functional forms and interpret the results as an approximation of the underlying joint envelope. Although appealing, this approach has a serious limitation: To use the expression in Ekeland et al. (2004), it does not "exploit" the model's economic content. In fact, the estimated hedonic price function may not be consistent with the assumptions about household demand in the second step. This problem arises because Rosen's "joint envelope" is not just an unspecified function that somehow falls between the bid-function envelope and the offer-function envelope, but is instead a function that must simultaneously equal both envelopes. Studies that estimate the hedonic price function using a general functional form cannot ensure that this condition is met and therefore cannot rule out the possibility of bias in the estimated coefficients.

One important innovation in the hedonics literature, for example, was the introduction of the Box-Cox technique by Goodman (1978) and Halvorsen and Pollowksi (1981). As it turns out, (18) converges to a Box-Cox form when σ_3 approaches infinity. In this case:[8]

$$\frac{\left(\hat{P}^E\{S_1,\ldots,S_n\}\right)^{\nu+1}-1}{\nu+1}=C_0^*+\sum_i\left(\frac{1-\sigma_{1i}}{\sigma_{2i}}\right)\left(\frac{\left(S_i\right)^{\zeta_i}-1}{\zeta_i}\right). \tag{22}$$

Standard Box-Cox procedures estimate a single value for ζ, which is equivalent to assuming that all amenities have the same price elasticity.[9] A procedure that uses implicit prices derived from a standard Box-Cox specification to estimate separate price elasticities for each amenity in the

Rosen second step is therefore inconsistent. In effect, this approach omits variation associated with differences in price elasticities across amenities and therefore could yield biased coefficients—and biased estimates of implicit prices.

Some recent studies estimate linear hedonic regressions. Bayer et al. (2007) justify this approach on the grounds that it yields the *mean* marginal willingness to pay, or mean bid function, for each amenity. In fact, however, the hedonic price function is an envelope, not a representative bid function, and as emphasized by Ekeland et al. (2004), bid functions and envelopes have different curvature. A simplified version of the Bayer et al. 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.[10] A bid function for a given household is found by solving for the price, say P^*, that holds utility constant (at V^*) over housing and neighborhood traits, or $P^*_{hi} = \left(\alpha_{0i} - V^*_i\right) / \alpha_{hi} + \sum_j \left(\alpha_{ji} / \alpha_{hi}\right) X_{hj}$. 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. The market price of housing is the envelope of these bid-functions.

Consider a single amenity, re-write the bid function as $P^*_i = a_i + b_i X$, and assume that the sorting equilibrium can be approximated 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. With a linear utility function, in other words, the hedonic equation is a quadratic function of the *X*'s.[11] By Bayer et al.'s own assumptions, therefore, the linear hedonic equation they estimate is mis-specified,[12] and their argument that the estimated coefficients indicate average preferences is not convincing.

Epple and Sieg (1999) and Epple et al. (2001) 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.[13] They estimate a bid-function envelope for a single amenity that 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 producing 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 also draws on the theory of bidding and sorting, but it 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 an unusual statistical technique.

Bayer et al. (2007) provide a clever alternative empirical approach to sorting based on discrete choice analysis.[14] They 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 households with different observable characteristics are willing to pay for each amenity. 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.[15]

Epple and Sieg (1999), Epple et al. (2001), Bayer et al. (2007), and Bajari and Kahn (2005) explicitly allow for heterogeneity in household preferences.[16] The approach developed here accounts for observed and unobserved heterogeneity in the demand for S and H in the first step (where it is integrated out) and through the error term in the second-step regression, (11).

Another important problem confronting any hedonic price regression is that omitted variables may lead to biased estimates for amenity variables. One approach to this problem is to take advantage of unusual arrangements in school district boundaries, as 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). Another approach is to collect time-series data on school-district performance so that time-invariant unobservables can be removed with district fixed effects (Clapp et al. 2008). A third strategy is to estimate the hedonic with instrumental variables. Good instruments for this approach are difficult to find. Policy choices by higher levels of government lead to plausible instruments in some studies, such as the Chay and Greenstone (2005) study of air pollution. In most other cases, such as Weimer and Wolkoff (2001) or Downes and Zabel (2002), the instruments are neighborhood characteristics, and it is difficult to see why they should be considered exogenous.

A fourth approach, which can be applied to variation in school quality within a school district, appears in Black (1999), Kane et al. (2006), and Bayer et al. (2007). The idea behind this approach is that houses near the same school attendance-zone boundary share many unobserved neighborhood traits. A set of fixed effects for the houses within a certain distance of each boundary segment is included in the hedonic regression to control for these traits and to eliminate any bias that would arise if they were left out of the regression. Observations far from attendance zone boundaries are dropped.[17]

Kane et al. (2006) and Bayer et al. (2007) point out that this approach needs to consider sorting. As Bayer et al. put it, "even if a school boundary was initially drawn such that the houses immediately on either side were identical, households with higher incomes and education levels might be expected to sort onto the side with the better school" (p. 590). This sorting causes a problem, because it leads to differences in the demographics of the neighborhoods on either side of a school boundary; if households care about the demographic composition of their neighborhood, these differences may lead to differences in housing prices. To address this problem, Bayer et al. include neighborhood demographic traits in the hedonic regression.

The analysis in this paper leads to a different conclusion. The key problem is that sorting results in different bids on either side of the border, as indicated by different values of ψ in (12) or (13). Consider (13) with one school variable, S, and one neighborhood variable, N, and suppose that N is shared but S differs within neighborhoods:

$$\begin{aligned}\ln\{P\} &= C + \psi_S\left(\frac{\mu_S}{1+\mu_S}\right)S^{(1+\mu_s)/\mu_s} + \psi_N\left(\frac{\mu_N}{1+\mu_N}\right)N^{(1+\mu_N)/\mu_N} \\ &= C + \psi_S\left(\frac{\mu_S}{1+\mu_S}\right)S^{(1+\mu_s)/\mu_s} + \bar{\psi}_N\left(\frac{\mu_N}{1+\mu_N}\right)N^{(1+\mu_N)/\mu_N} + \left(\psi_N - \bar{\psi}_N\right)\left(\frac{\mu_N}{1+\mu_N}\right)N^{(1+\mu_N)/\mu_N} \quad .\end{aligned} \tag{23}$$

Thanks to sorting, the households on one side of the border differ from those on the other side and therefore have different values of ψ_N. The neighborhood fixed effects remove the average price impact in each neighborhood, indicated by the first term containing $\bar{\psi}_N$, but ignore deviations around this effect associated with sorting, indicated by the last term. Because the average effect is shared by households on both sides of the border, it is not highly correlated with

S and is therefore not an important source of bias. In contrast, the last term is likely to be correlated with *S*, and hence a source of bias, because it reflects household traits associated with sorting. In short, the main problem with the attendance zone boundary approach is that sorting leads to a situation in which the households on either side of a boundary place a different value on the shared characteristics of that boundary neighborhood. This problem calls for interaction terms between demand determinants and amenities (or a procedure to integrate out these determinants), not a set of neighborhood demographic characteristics.

It might also be true that households care about neighborhood demographic traits such as average income and average education, which are two of the variables in the Bayer et al. regressions. Income and education are also demand variables that influence willingness to pay for amenities, however, and the Bayer et al. procedure does not separate these two possibilities. To the extent that income and education pick up variation in demand, the Bayer et al. approach is analogous to (12), which describes the family of bid functions, not to (17), which describes the envelope. This distinction is crucial because the *N* variables, such as income and education, in (12) pick up variation in bids across household types, not variation in bids based on neighborhood amenities. Moreover, (12), unlike (17), is subject to bias from unobserved demand determinants.[18] Overall, the ambiguity that arises in interpreting the Bayer et al. equations casts doubt on their claim that estimates of MWTP for school quality are overstated if neighborhood income and education are not included in the hedonic price regression.[19]

This paper recognizes that some neighborhood demographic traits may serve as amenities that directly affect housing bids. To minimize the above problems of interpretation, I distinguish demographic traits, such as ethnicity, that can be readily perceived by households from demographic traits, such as average income or education, that are determinants of amenity demand and that are difficult for households to perceive. Then I assume that demographic traits in the first category are neighborhood amenities that directly affect housing bids whereas those in the second category serve only as demand determinants, not as amenities. Although this assumption cannot be directly tested, this paper provides indirect evidence by determining whether variables in the second category affect household demand for amenities. In addition, I minimize bias from omitted amenity variables by including in the hedonic price regression extensive information on non-demographic neighborhood amenities.

As noted earlier, endogeneity is the principal challenge facing the second step is in the Rosen procedure. Many studies estimate the second step with instrumental variables techniques, but good instruments are difficult to find (Sheppard 1999 Taylor 2008). Ekeland et al. (2004) estimate the first stage hedonic with non-parametric methods and then identify the second-stage equation by exploiting the difference in the curvatures of the envelope and bid functions. Bajari and Kahn (2005) assume a unitary price elasticity of demand for S. In this case, substituting (8) into (3) and multiplying by S leads to a new regression with the implicit price multiplied by S as the dependent variable and only exogenous variables on the right side. This approach cannot be generalized, however, to non-unitary price elasticities. The approach proposed here avoids this endogeneity problem by estimating μ directly, with no need for a two-step procedure. Moreover, the other demand parameters for S and H can be estimated via (11), which does not have S on the right side for any value of μ.

Estimation Procedures

Equation (21) provides a one-step method for obtaining 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 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, 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, controlling for other geographic variables.[20]

Data

The data used for this study build on the data set described in Brasington (2007) and Brasington and Haurin (2006), which consists of all the house sales in Ohio in 2000. This study makes use of the sub-sample for the Cleveland metropolitan area. As discussed more fully below, the data set contains information on sales price, housing characteristics, housing location, census block group characteristics, school performance, crime rate, and air quality. These data were supplemented with various 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 neighborhood fixed effects. I use census block groups (CBGs) as neighborhood units, although I break up the few 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.

The second challenge is to account for locational characteristics that vary within neighborhoods. Seven distance-based variables vary across houses 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, to the Cleveland airport, and to the center of the CBG. The signs of the last two variables are not clear a priori. Houses close to an airport might be more valuable due to the convenient access or less valuable due to the noise. Houses located on the outskirts of an established neighborhood might be less valuable, but it is not clear whether CBGs accurately identify established neighborhoods.

Because the hedonic equation includes CBG fixed effects, these variables are specified to measure only within-CBG variation. Except for the first variable, which has a specification based on urban models, these variables are specified as the difference between the log of the distance to the house and the log of the distance to the center of the CBG. This approach is equivalent to including distances directly in a Cobb-Douglas utility function. The first-stage hedonic also controls for differences between a house and the center of the CBG in the test scores of the nearest elementary school or in location within a historic district. These differences equal zero for most observations, but not for all.

The worksite variables come from a basic urban model. 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. The center of each worksite was defined as the employment-weighted centroid of the zip codes in the relevant cluster. Following Wieand (1987) and Yinger (1992), the boundary between the residential zones of workers commuting to two different worksite was set where the bid functions (based on Cobb-Douglas utility functions and assumed commuting costs) to each worksite were equal. Final boundaries were set so that assigned jobs had the same distribution across worksites as actual jobs. These procedures lead to one commuting variable for each worksite, defined as the bid function for the house’s location divided by to the bid function for the CBG’s location. The coefficient of each variable logged is the inverse of the share of income spent on housing. See Yinger (2009).

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 22,880 observations. The R-squared is 0.7893. The housing characteristics are, with the exception of "One Story," highly significant with the expected sign. The set of 1,665 geographic fixed effects is also highly significant. The commuting variables are all significant at the 0.1 percent level. Their coefficients are all close to 3.4, which has the reasonable implication that the share of income spent on housing is about 29 percent. Additional significant impacts on housing prices come from the distance from an environmental hazard (positive), from Lake Erie (negative), and from the center of the CBG (negative). The latter result suggests that households are willing to pay for a location at the center of an established neighborhood. In contrast to the results in Kane et al. (2007), these results do not indicate that households are willing to pay for access to elementary schools, at least not in the distance range in these data.

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, 1,665, 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 of full time workers in the CBG 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.[21] 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 or if commuting speeds differ by direction.[22]

The second step is to select service and "amenity" variables. As indicated in Table 3, I selected seven of these variables: the test scores of the neighborhood elementary school, test scores in the school district, a high school quality indicator, safety, distance to the nearest environmental hazard, and two measures of ethnic composition. The test score variables measure the passing rate on all five examinations (mathematics, reading, writing, science, and citizenship) that form the basis of the state's accountability system.[23] These scores are averaged over the school years 1997-98 and 1998-99. The elementary variable, which refers to fourth grade scores, is the score in the nearest elementary school in the neighborhood's school district relative to the average score in the district. The district test score variable is the passing rate on the state tests in ninth grade. 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). The safety measure is the inverse of the crime rate, which equals total crimes per thousand people in 1997. The ethnic composition measures are the inverse of the percent black and of the percent Hispanic in the CBG.

Table 3 also lists tax variables. A few school districts in the Cleveland area levy an income tax, usually at a 1 percent rate. Ohio requires assessment at market value so the nominal property rates are roughly equivalent to effective rates.[24] 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 neighborhood controls. These variables could 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 unsuitable for the functional forms derived earlier. These variables, listed in Table 4, include worksite and county fixed effects, as well as indicators of location near a wide range of factors that might influence housing prices.[25]

A few variables require special comment. The eastern suburbs of Cleveland receive unusual amounts of snow. The snow-belt variables 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. 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 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 from each CBG to the nearest cluster. The final variables measure whether a CBG was within 20 miles of the nearest cluster and the distance from the cluster if it was less than 20 miles. These air-pollution variables were interacted with direction to account for the possibility that air pollution is affected by the wind.

These steps make it possible to estimate the full bid function envelope, (21) without $H\{X\}$. This equation is difficult to estimate in its full non-linear glory, so I assume that $v = -1.0$. As shown by (13), this leads to $\log\{V\}$ as the dependent variable with no $(1+v)$ terms on the right side. Second, I assume that $\gamma = 1.0$, which simplifies the commuting variables. These assumed elasticities are somewhat larger in absolute value than most estimates (Zabel 2004).[26]

The model was estimated with nonlinear least squares (using the STATA "nl" procedure with the "hc3" option to account for heterockedasticity). Starting values were found through a grid-search procedure that identified the lowest sum of squared errors (SSE) for various values of the seven price elasticities, one for each service or "amenity."[27] For the Elementary School variable, SSE continued to decline as μ increased in absolute value, so I set this elasticity at minus infinity, which is equivalent to setting the first exponent in (17) to 1 and the second to $(1+(1/\sigma_3))$.[28] I estimated the model with three values for σ_3—½, 1, and 2—to allow for negative, zero, and positive curvature, respectively, in the $S\{\psi\}$ function. The first of these options leads to the lowest SSE. I also discovered that the SSE can be lowered further by setting $\sigma_3 = 2$ for the High School variable, while keeping it at ½ for every other variable; as a result, this set of values forms the basis for my preferred estimations.[29]

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 a standard specification in which the amenity variables enter linearly. All of these variables except Non Black are statistically significant with the expected sign. In the second column, which sets all the price elasticities to -1, both terms in (17) are significant for four of the variables, and one term is significant for two others. Neither term for the safety variable is significant, however.

The last three columns show results for the full model with various values for σ_3; each of these regressions has lower SSE than the simpler models. My preferred estimates are in column three.

Envelopes

All the amenity elasticities in column 3 (except the one for Elementary Scores, which, as indicated earlier, is set to $-\infty$) are significant and less than 1.0 in absolute value. The elasticity is -0.154 for 9th Grade Score, -0.655 for High School, -0.284 for Safety, -0.228 Distance to Hazard, -0.250 for Non Black, and -0.239 for Non Hispanic.[30] The estimated value for σ_1 also is statistically significant in every case. As discussed earlier, the sign of σ_2 provides a test 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 σ_2 coefficients are all positive, and they are significant for four of the variables (Elementary School, High School, Non Black, and Non Hispanic). These results therefore support the sorting hypothesis in these four 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 10 to 20 percent.[31] 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 3 to 5 plot the estimated envelopes and illustrative underlying bid functions.[32] Figure 3 is based on the school variables. The envelope for Elementary Score is relatively flat until a school's score exceeds the district average by about one-third. Then the envelope slopes upward, and housing prices at the maximum value of the variable (3.0) exceed the minimum price by 20.1 percent. Because of the infinite price elasticity, (13) implies that the bid functions are straight lines. The envelope for 9th Grade Score indicates that school districts with the lowest scores have much lower prices than other districts, but the impact of scores on housing prices levels off once an intermediate score has been reached. More specifically, moving from the minimum to the mean score raises housing prices by 73.1 percent, whereas the impact of moving from the mean to the maximum score is only 2.6 percent. These differences are more dramatic along underlying bid functions. For the bid function on the left, moving from the minimum to the mean score raises the bid by 348.6 percent; moving to the maximum adds only 0.2 percent.

The pattern for High School is quite different. Raising High School 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 24.3 percent. Below the median value, however, this envelope has a negative slope. Going from the minimum to the median value would lower the envelope by 34.9 percent. This result demonstrates the importance of accounting for household heterogeneity. As illustrated by the upward-sloping bid function, some households value the school district quality dimensions captured by High School; these households would pay 41.2 percent more to go from the minimum to the median value of High School and they would pay 9.0 percent more to go from the median to the maximum. In contrast, other households either dislike these dimensions of district quality or else prefer other, unobserved dimensions that are negatively correlated with them. The illustrative downward-sloping bid function in this panel, for example, applies to a household that would pay 29.2 percent less to go from the minimum to the median value of High School and 8.2 percent less to go from the median to the maximum. One possibility is that these households value 10^{th} through 12^{th} grade outcomes for students in the middle of the distribution, which are likely to be negatively correlated with the extremes as measured by High School.

A careful look at Figure 3 indicates that the envelopes for Elementary Score and 9^{th} Grade Score also have segments with negative slopes. The segment for 9^{th} Grade Score involves only 31 observations and does not appear to raise substantive issues, but the segment for Elementary scores is quite long, even though it is much flatter than the segment for High School. These segments provide valuable insight into household heterogeneity, and the determinants of these negative slopes are explored in the following section.

Figure 4 applies to Safety and Distance from Hazard. The envelope for Safety is similar to the one for High School. This figure is plotted with the crime rate (the inverse of Safety) on the horizontal axis. Housing prices decline 17.1 percent as one moves from the lowest crime rate up to about twice the mean crime rate (60.8), but then increase 21.5 percent from this point up to the maximum crime rate. The illustrated bid functions indicate a large willingness to pay (69.3 percent) for the lowest- versus the highest-crime neighborhoods among some households and a modest willingness to pay (17.9 percent) for the opposite difference among others. The negatively sloped part of the envelope indicates that some households value unobserved neighborhood traits that are positively correlated with the crime rate—a result explored below. The envelope for Distance to Hazard is similar to the one for 9^{th} Grade Scores, although the magnitudes are smaller. Housing prices increase 14.8 percent as one moves from next to a hazard to one mile away, but then increase only 0.7 percent more as one moves to the maximum distance in the sample. The comparable changes on the illustrative bid function are 169.0 percent and 0.2 percent. The negatively sloped segment of this envelope applies to only 7 observations.

Envelopes for the ethnicity variables are drawn in Figure 5. The first panel shows that the bid envelope declines with percent black up to 75 percent but increases thereafter. More specifically, going from zero to 75 percent black lowers the envelope by 26.2 percent, whereas going from 75 to 100 percent black raises the envelope by 4.5 percent. These results support the view that households have heterogeneous preferences, as 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. The second panel reveals a strikingly similar pattern for neighborhood Hispanic composition. Moving from zero to 49 percent Hispanic lowers the envelope by 42.6 percent, but moving from 49 to 63 percent Hispanic (the maximum composition in the sample) raises the envelope by 6.7 percent. These results indicate heterogeneity in preferences for Hispanic composition, too.

Table 6 presents results for geographic controls. Housing prices are significantly higher within two miles of Lake Erie. This effect declines with distance from the lake, although the phase-out is significant at only the 10 percent level. The coefficients of the two snowbelt variables imply a maximum positive impact on housing prices at 9.8 miles from Lake Erie in the eastern suburbs, which is exactly where snowfall is the greatest. Households apparently value snowfall or some topographical feature with which it is correlated. Location in or near the black ghetto also lower housing prices by 9.2 and 8.1 percent, respectively, controlling for ethnic concentration. The first result suggests that blacks are concentrated in neighborhoods with unobserved disamenities, perhaps because of past discrimination. The second result could be a sign of unobserved disamenities in neighborhoods bordering the ghetto or of residents' concerns about future neighborhood change.

The two airport variables (the first of which is significant at only the 10 percent level) indicate that housing prices are about 4 percent higher in the neighborhoods closest to the airport, which are about 2 miles from the airport's center, and then decline with distance from the airport. In combination, these results, which might reflect a trade-off between the value of access to the airport and its jobs versus the disadvantages of airport noise, indicate that housing prices are higher within 5 miles of the airport but lower 5 to 10 miles from the airport. Housing prices also decline with the dumping of pollutants onto land in the neighborhood and are lower in locations within 20 miles of a major air pollutions site, an effect that declines with distance from the site. The air quality effects do not arise to the northwest of these sites, probably because two sites are located near Lake Erie with few, if any, houses to the northwest, and the third is located to the southeast of another site, so that moving to the northwest does not imply cleaner air. Finally, housing prices are boosted by neighborhood parks but 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 (15), and then to estimate (11).[33] This step is complicated by the negative values of ψ for some observations. In three cases (9th Grade Score, Distance to Hazard, and Non Hispanic), the number of observations with negative ψ is small, so these observations are simply dropped. In the other cases, we can account for the large number of negative ψs by adding a multiplicative term, Φ, to the right side of (8), and hence to (11), that determines the sign of MB_S as a function of N, Y, and Q, where Q is a set of demand variables for amenities that are correlated with S. This function, $\Phi = \Phi\{N, Y, Q\}$, distinguishes households that value increments in S more than they value increments in correlated (but unobserved) amenities ($\Phi = 1$) from households with the opposite preferences ($\Phi = -1$).

With this extension, (11) can be estimated using an endogenous switching regression (Lokshin and Sajaia 2004).[34] The "switch" is between positive and negative values for ψ. This approach simultaneously estimates $\ln(\psi)$ as a function of household traits N, M, and Y (at the CBG level), when ψ is positive; $\ln(-\psi)$ as a function of N, M, and Y when ψ is negative; and the sign of ψ (1=negative; 0=positive) as the realization (Φ) of a latent variable that depends on N, Y, and Q. This model reveals which traits are associated with a negative ψ (that is, with a negative MR_S), and then estimates (11) separately for households with positive and negative ψ.

Elementary Score requires special treatment. First, as shown in Figure 3, the portion of the envelope with a negative slope is almost flat; with so little variation in the slope, the determinants of the slope cannot be estimated. In this case, therefore, I estimate (using the STATA "heckman" command) a selection equation that determines whether the slope is positive combined with (11) for positive values of ψ. Second, because $\mu = -\infty$, the terms from the service demand equation except for Φ drop out of (11). In this case, normal sorting never occurs and the exponent of the income term is $-\gamma$. As shown in the first column of Table 7, the estimates of (11) for Elementary Score are consistent with this prediction. The income coefficient is negative and significant. Its absolute value, 0.84, is somewhat higher than recent estimates of γ (Zabel 2004), but close to the value I assumed for the commuting variables.

Tests of normal sorting for the other amenities appear in the remaining columns of Table 7. Consider first the results for the positively sloped portions of the envelopes. Normal sorting requires $(\theta/\mu + \gamma) < 0$, which is the same as a positive value for the income exponent in (11). Estimates of this exponent are positive and significant for 9th Grade Score, High School, Safety, and Distance to Hazard, indicating normal sorting in these cases. The results for Non Black also support normal sorting where this trait is regarded as an amenity, but the income term is not significant for Non Hispanic. For the negatively sloped portions of the envelopes, a negative sign for the income term supports normal sorting. Hence, the result for High School supports normal sorting, but the result for Safety does not. The Wald tests indicate that the demand parameters with $\Phi = 1$ are significantly different from those with $\Phi = -1$ in all cases except Non Black.

Table 8 presents representative selection equations, Φ, for Elementary Score, High School, Safety, and Non Black. For Elementary Score, the coefficients indicate the impact of a household characteristic (a CBG average) on selection into the *positively* sloped segment of the envelope. In other words, they identify household traits associated with a willingness to pay more pay more for housing based on the relative quality of the nearest elementary school. As one would expect, married households and households with more education are more likely to be on this segment. This likelihood decreases with the share of households affiliated with a Christian church, which may indicate a preference for a religious school, but increases with share of children in private school, which may indicate neighborhoods in which people place a high value on education. The school-district context also matters: Location on the positively sloped segment is more likely in high-welfare districts, but less likely in high-minority districts.[35]

The coefficients for the other three selection equations indicate the impact of a household trait on the likelihood of being on the *negatively* sloped portion of the envelope. The probability of a negative ψ for High School is clearly linked to variables that indicate a low demand for public education. To be specific, this probability decreases with the number of households with children and with household education, and increases with unemployment and with the share of households sending their children to private schools.[36] In the case of Safety, the probability of a negative ψ increases with the unemployment rate and decreases with the share of the population that is married, has advanced education, has children, or is elderly. These results are consistent with the view that families and the elderly place a particularly value on safety, but that single people and people who are unemployed have different priorities.[37]

In the last column of Table 7, which applies to Non Black, a negative ψ indicates a preference for a largely black neighborhood. These results are not as illuminating as those in other columns, because sorting in this case obviously depends heavily on the ethnicity of a household, which, for reasons explained earlier, cannot be included directly in the analysis. As a result, the results largely capture the correlation between percent black and other household traits. More specifically, a negative ψ is less likely for households that have more education, that are more likely to be employed, that are married, that have a blue collar job, or that have children in private school. The last two variables also might be indicators of white prejudice. Many largely black neighborhoods in Cleveland also have high homeownership rates and low turnover, and the probability of a negative ψ also increases with these traits.

These results indicate that envelope-based estimates of household heterogeneity in amenity demand follow reasonable patterns. Because they show that income and education are key determinants of this demand, these results also highlight the potential for endogeneity bias when income and education are treated as neighborhood amenities, as in Bayer et al. (2007). A key challenge for future research is to determine whether these finding about household heterogeneity are robust to more general specifications of the amenity demand functions and the $S\{\psi\}$ approximation.

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 neighborhood amenities and housing and the theory of household sorting, I derive the bid function envelope that arises when households have heterogeneous preferences for amenities and housing. 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 amenities and housing. Second, it yields estimates of the price elasticities of demand for 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 conduct 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 the Cleveland area in 2000. The regressions, which control for a wide range of housing and neighborhood characteristics, indicate a strong impact on housing prices from seven different public services or "amenities": elementary test scores, ninth grade test scores, high school quality, safety, distance from an environmental hazard, percent non black, and percent non Hispanic. Preliminary tests led me to set the price elasticity of demand for Elementary Score at $-\infty$, but price elasticities ranging from -0.15 to -0.66, all precisely estimated, were obtained for the other six variables.

The estimated bid envelopes and underlying bid functions exhibit several different shapes. For 9th Grade Score and Distance to Hazard, moving from the lowest level of the 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 variables, the impact of an amenity increase on housing prices does not fade out at high levels of the 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 four "amenities" (Elementary Score, High School, Non Black, and Non Hispanic) 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 usually "normal," with higher-income households sorting into locations where the services and amenities are more desirable.

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.

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. 1987. “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.

Cheshire, Paul, and S. Sheppard, 2004. "Capitalising the Value of Free Schools: The Impact of Supply Characteristics and Uncertainty." *Economic Journal* 114: F397–F424.

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.

Figlio, David N., and Maurice E. Lucas. 2004. "What's in a Grade? School Report Cards and the Housing Market." *The American Economic Review* 94 (3) (June): 591-604.

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.

Goodman, Allen C. 1978. "Hedonic Prices, Price Indices and Housing Markets." *Journal of Urban Economics* 5 (October): 471-484.

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.

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.

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.

Rouwendal, Jan. 1992. "The Hedonic Price Function as an Envelope of Bid Functions: An Exercise in Applied Economic Theory." *Journal of Housing and the Built Environment* 7 (1) (March): 59-80.

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.

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.

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.

Yinger, John. 2009. "The Price of Access to Jobs: Bid Function Envelopes for Commuting Costs." Unpublished Manuscript, Syracuse University, May.

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] Rouwendal (1992) provides an alternative envelope derivation with a linear utility function. He specifies *C* as a quadratic function of observable *N* variables and then solves for the slope of the envelope. Rouwendal also makes the same claim that I do, namely, that "The second stage of Rosen's procedure is circumvented in this way and so are the problems associated with it" (p. 59). This claim is not correct, however, because his approach neither integrates out unobservables nor leads to an estimating equation in which demand parameters are identified. Moreover, Rouwendal estimates the analog to (12), not the envelope.

[4] Yinger (1982) 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).

[5] An equivalent assumption is that household demand for S_i reflects the systematic determinants of S_j but not the unobserved and random factors.

[6] 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).

[7] Price discrimination against certain groups also 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).

[8] The constant term now has an asterisk to indicate that it has been adjusted to account for the addition of the 1s in the numerators of the Box-Cox expressions.

[9] This form is usually applied to all neighborhood and housing characteristics, so it is also restricting neighborhood and housing characteristics to all have the same price elasticity.

[10] This formulation leaves off several terms in the Bayer et al. formulation, their equation (2); this simplification facilitates the exposition but do not alter the conclusions in the text.

[11] This is the envelope derived by Rouwendal (1992). The quadratic utility function in Epple (1987) and Ekeland et al. (2004) also leads to a quadratic envelope, but it has interactions between services/amenities and a different specification of the preference parameters.

[12] Bajari and Kahn (2005) use a utility function that 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 $P_i= (Y_i-Ui^0)+b_i(\ln\{X\})$. 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.

[13] 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.

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

[15] As implemented, their approach also has the weakness that it employs their mis-specified hedonic price equation (discussed elsewhere in this section) in a first stage.

[16] 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 similar, which is consistent with sorting, that is, with equal household unobservables within neighborhoods.

[17] Dropping observations far from boundaries not only lowers precision but may also introduce bias if households recognize the possibility of boundary changes and therefore place less weight on school quality differences when they live near a border (Cheshire and Sheppard 2004).

[18] Brasington and Hite (2008) address this type of endogeneity with a simultaneous equations procedure. They add individual household traits (obtained through surveys) to the Ohio data on which this paper draws and estimate an analog to (12) along with equations to determine key household traits. Unfortunately, however, the instruments available for the household-trait equations are also demand determinants and may not be exogenous.

[19] Bayer et al. also include two other explanatory variables that are likely to be endogenous: density and land use. Urban models show that density is a function of housing price (Brueckner 1987), and residential uses dominate in locations where households bid a lot for housing.

[20] Borjas (1987) shows that the standard errors for the second-stage coefficients need to be corrected for heteroskedasticity.

[21] The details of these procedures are presented in Yinger (2009).

[22] Bajari and Kahn (2005) and Bayer et al. (2007) use unadjusted time to work, which introduces endogeneity bias if unobserved determinants of job location are correlated with housing bids.

[23] Preliminary tests yielded similar results with test scores averaged over grades 3, 6, and 9 or 6 and 9. The data set includes 4th grade scores in 1997-98 and 6th grade scores in 1999-2000, which yield a value added measure (except for movement in and out of the district). A preliminary examination of this variable duplicated the finding in Brasington and Haurin (2006) and Downes and Zabel (2002) that a value-added measure does not affect house values.

[24] Based on assessment-sales ratios for 2005 to 2008, assessments appear to be reasonably accurate: the 2005 ratios averaged 82.2 percent, were highly correlated across time, and 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. These data come from http://tax.ohio.gov/channels/research/property_tax_statistics.stm.

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

[26] Although I do not estimate these elasticities, some evidence on the plausibility of my assumptions is presented in a later section. Estimates of γ, which require more complicated functional forms for the commuting variables, are provided in Yinger (2009).

[27] Because ν is assumed to equal -1, this grid search can estimate tax capitalization linearly by making use of the approximation that $\ln\{r+\beta\tau\} \approx \ln\{r\} + (\beta/r)\tau$.

[28] As shown earlier, these bid functions apply to Bayer et al. (2007), which implies that they implicitly assume infinite price elasticities for all housing and neighborhood traits. This envelope was derived by Rouwendal (1992), but he does not mention price elasticities or ψ.

[29] Given the complexity of this model, a grid search to find the SSE-minimizing values of σ_3 for each service/amenity was not feasible. However, I did estimate the model with integer values of σ_3 from 3 to 10 and found that raising σ_3 inevitably raised the SSE. Estimating the model with fractional values of σ_3 below ½ is also not feasible because the exponent of the second term in (17) becomes positive, and the term becomes very large, when $\sigma_3 < -\mu/(1+\mu)$. This feature of the model is what led me to try a higher value of σ_3 for the high school variable.

[30] Using a modified median voter model and school district data, Duncombe and Yinger (1998) estimate a price elasticity equal to -0.313 for an aggregate test score measure in New York State.

[31] Correcting for the average assessment-sales ratio in 2005 (see note 25) 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 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.

[32] These figures are based on Table 5, column 3; figures for columns 4 and 5 are similar.

[33] I estimated (21) with a Box-Cox form on the left side (defined as rent using $r = 0.03$ and the tax terms with their estimated coefficients in the capitalization rate) and right side variables defined by the elasticities in the text. The estimated Box-Cox parameter, $(1+v)$ in this case, is -0.0035 (not significantly different from zero), which indicates $v = -1.0035$. This informal estimate suggests that $v = -1$ may be a reasonable approximation.

[34] The endogenous switching regressions were estimated with the STATA "movestay" command with robust standard errors; an appendix with the full results is available from the author.

[35] Selection onto the positive segment is also less likely if a higher share of households have children, are foreign born, or live in a rural area and is less likely in locations with more educational homogeneity.

[36] The other results are consistent with this interpretation, too. Parents who speak English as a second language may be particularly concerned about public education, households affiliated with a Catholic or other Christian church may support public education less even controlling for private school attendance, and support for public schools may increase with long-term residence.

[37] Two other results in the safety equation do not have a clear interpretation: The probability of a negative ψ is higher for Asians and in neighborhoods with a high share of homeowners.

Figure 1: Deriving the Bid-Function Envelope

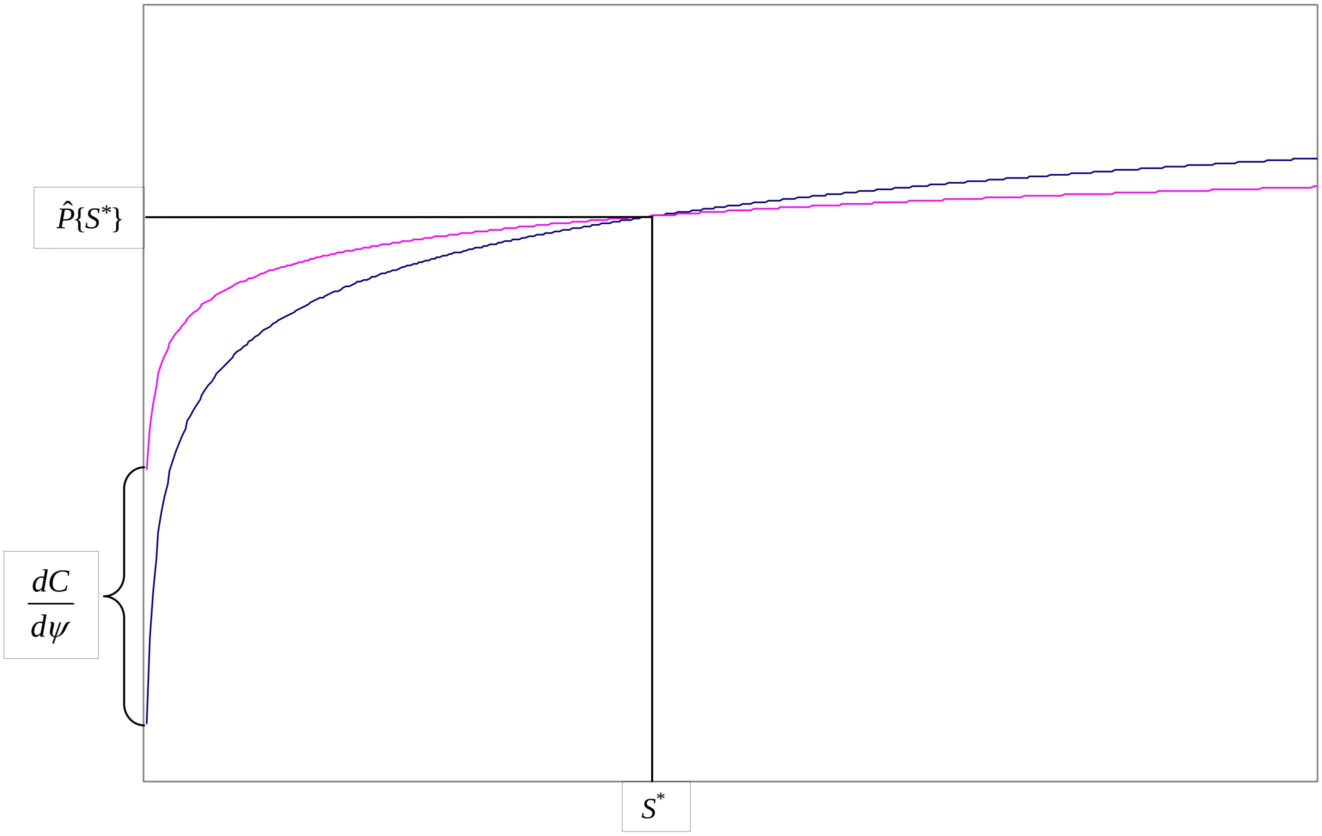

Figure 2. Supply and Demand for School Quality

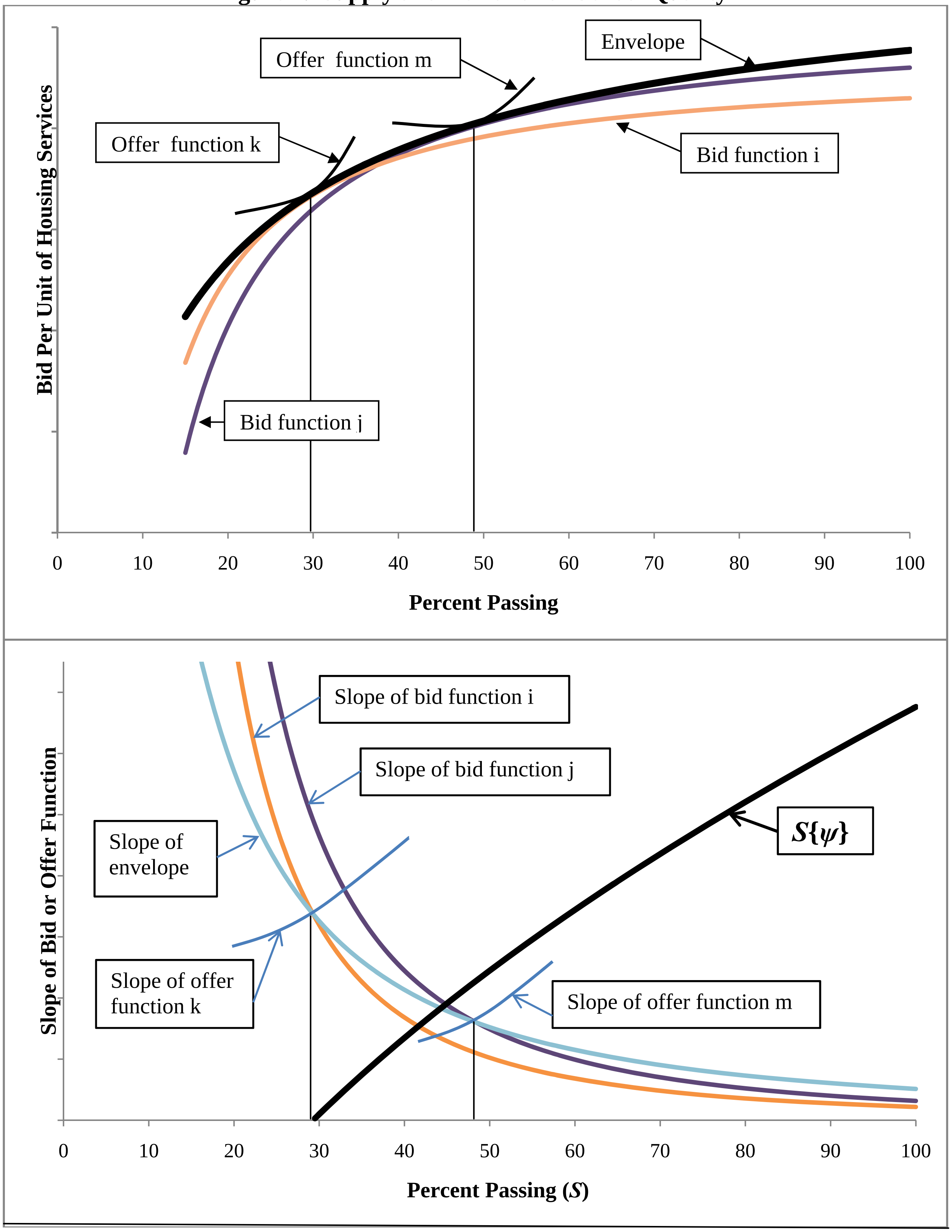

Figure 3: Bid Function Envelopes for School Variables

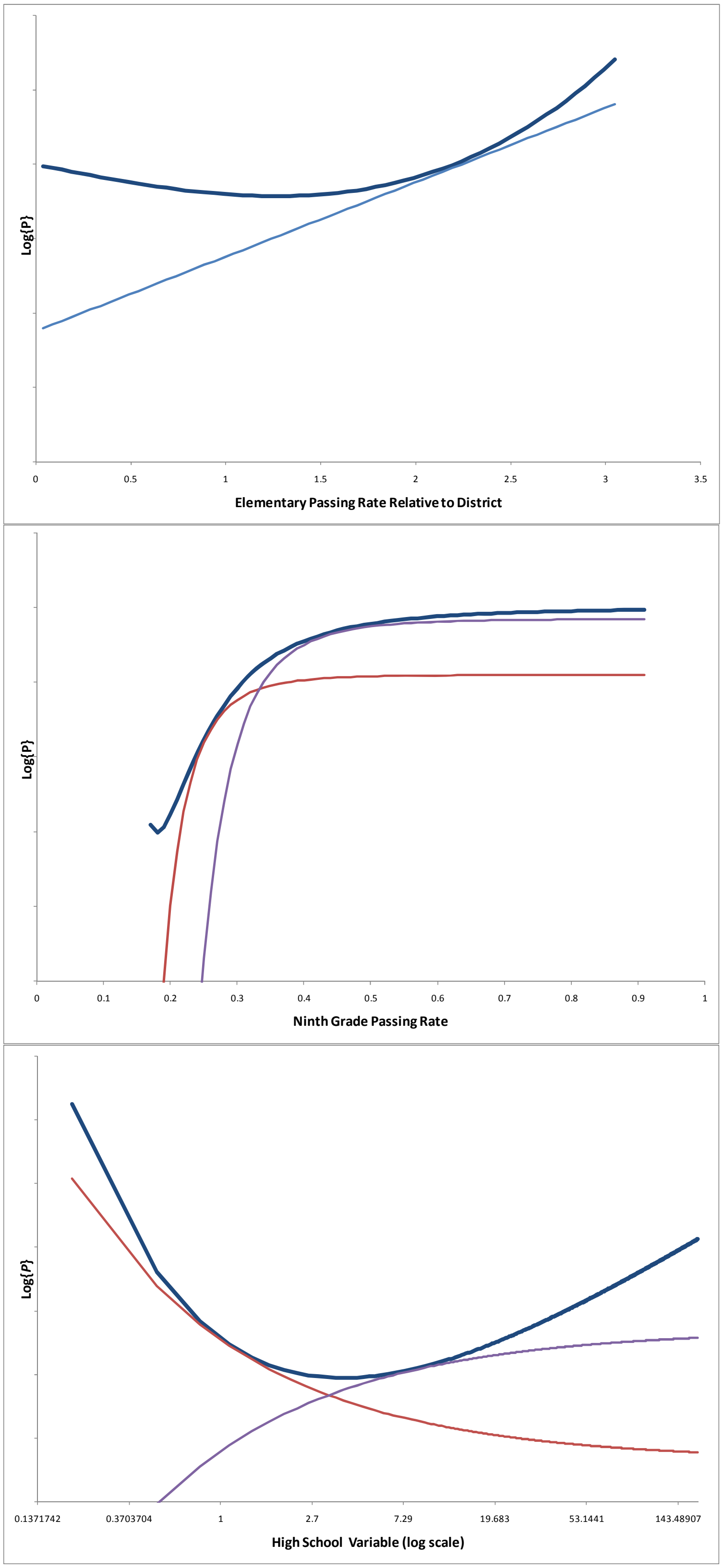

Figure 4: Bid Function Envelopes for Safety and Distance to Hazard

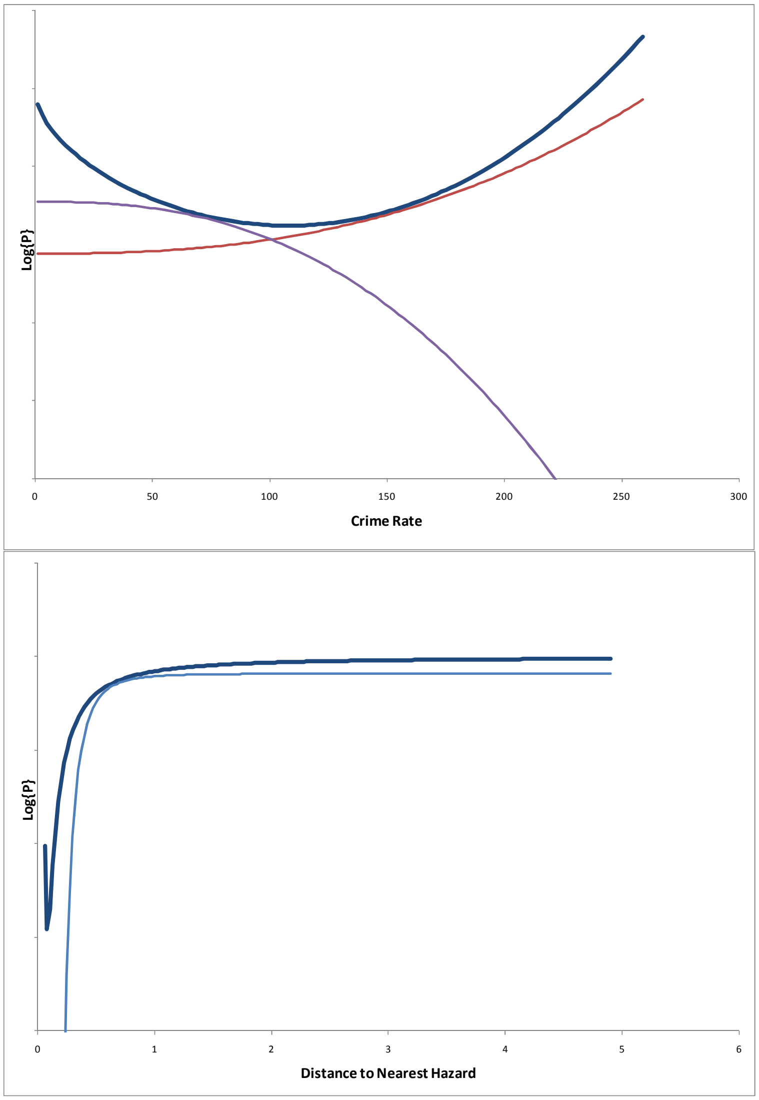

Figure 5. Bid Function Envelopes for Ethnicity Variables

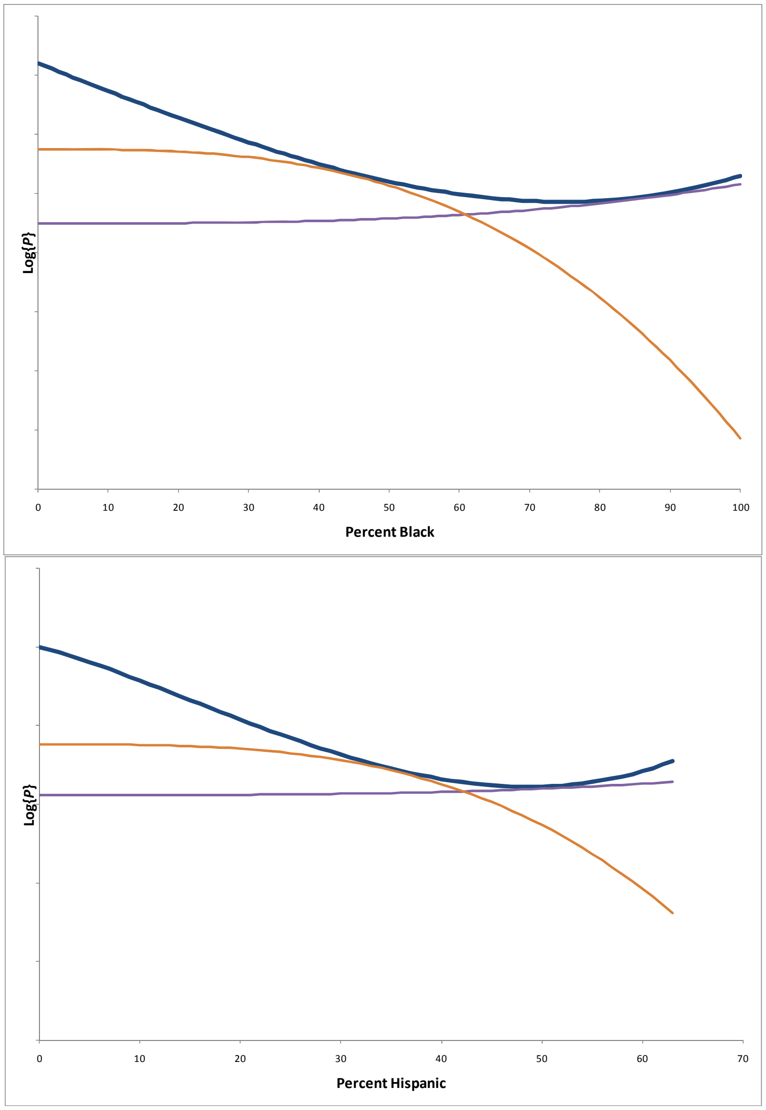

Table 1. Variable Definitions for Basic Hedonic with Neighborhood Fixed Effects

Variable	**Definition**
One Story	House has one story
Brick	House is made of bricks
Basement	House has a finished basement
Garage	House has a garage
Air Cond.	House has central air conditioning
Fireplaces	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
Outbuildings	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
Dist. to Pub. School[a]	Distance to nearest public elementary school in district (house minus CBG)
Elem. School Score[a]	Average Math and English test scores of nearest public elementary school relative to district (house minus CBG)
Dist. to Pri. 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 2; 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)
Dist. to CBG Center	Distance from house to center of CBG
Historic District[a]	Location is in historic district on national register (house minus CBG)

Note: Distances are all measured in miles.
[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	-0.00718	0.00496	-1.45	0.148
Brick	0.01568	0.00522	3.00	0.003
Basement	0.03076	0.00505	6.09	0.000
Garage	0.14023	0.00674	20.82	0.000
Air Cond.	0.02589	0.00552	4.69	0.000
Fireplaces	0.03193	0.00377	8.47	0.000
Bedrooms	-0.00850	0.00285	-2.98	0.003
Full Baths	0.06030	0.00425	14.18	0.000
Part Baths	0.04144	0.00413	10.03	0.000
Age of House	-0.08326	0.00324	-25.73	0.000
House Area	0.42435	0.00860	49.37	0.000
Lot Area	0.08457	0.00369	22.95	0.000
Outbuildings	0.12811	0.03966	3.23	0.001
Porch	0.03341	0.00726	4.60	0.000
Deck	0.05442	0.00526	10.35	0.000
Pool	0.09150	0.01798	5.09	0.000
Date of Sale	0.00017	1.81E-05	9.59	0.000
Commute 1	3.29460	0.67830	4.86	0.000
Commute 2	3.40253	0.70164	4.85	0.000
Commute 3	3.41496	0.70608	4.84	0.000
Commute 4	3.54144	0.73192	4.84	0.000
Commute 5	3.46484	0.71259	4.86	0.000
Dist. to Pub School[a]	0.00126	0.00451	0.28	0.780
Elem. School Score [a]	0.04079	0.02561	1.59	0.111
Dist. to Pri. School	-0.01341	0.01039	-1.29	0.197
Distance to Hazard	0.02339	0.00591	3.96	0.000
Distance to Erie[a]	-0.01261	0.00479	-2.63	0.009
Distance to Ghetto[a]	-0.13448	0.09462	-1.42	0.155
Distance to Airport[a]	0.14366	0.06236	2.30	0.021
Dis. to CBG Center	-0.01101	0.00287	-3.83	0.000
Historic District[a]	0.01708	0.01777	0.96	0.337
Constant	11.37008	0.01765	644.40	0.000

Notes: Dependent variable = log of transaction amount; 22,880 observations; R^2 = .7893; F(31, 21184) = 414.84 (significant at 0.000 level); 1,665 fixed effects with F(1664, 21,184) = 8.528 (significant at 0.000 level); estimated with "areg" command in STATA.

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

Variable	Definition
Elementary Score[a]	Average percent passing in 4th grade in nearest elementary school within school district on 5 state-specified tests (math, reading, writing, science, and citizenship) relative to the district average in 1998-99 and 1999-2000.
9th Grade Score	Average percent passing in school district in 9th grade on the 5 state tests 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 0.001 added to the denominator to account for CBGs with no blacks).
Nonhisp	Inverse of percent Hispanic in CBG (with 0.001 added to the denominator to account for CBGs with no Hispanics).
Commute	Urban model commuting variable (based on 1 minus estimated commuting cost to that location as fraction of income).
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).

Note: 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 (17) 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 2 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 20 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 center 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).
Historic District	CBG located in a historic district on the National Register of Historic Places

Note: 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	Estimated μ's σ_3 = 1/2[b]	σ_3 =1	σ_3 =2
Elementary Score 1	0.05425***	-0.05983	1.61031***	1.11481***	1.01265***
	(5.55)	(-1.79)	(4.29)	(7.61)	(14.10)
Elementary Score 2		0.20093**	32.12325***	12.02633***	5.70832***
		(2.65)	(4.78)	(3.96)	(3.55)
Elementary Scores μ			∞	∞	∞
9th Grade Score 1	0.56833 ***	0.26845*	0.03236***	0.18186***	0.42577***
	(17.11)	(2.07)	(14.35)	(57.04)	(77.83)
9th Grade Score 2		-0.30584	68.15567	159.6933	79.25396
		(-0.76)	(1.24)	(1.04)	(1.09)
9th Grade Score μ			-0.15433***	-0.14858***	-0.16617***
			(-8.25)	(-8.22)	(-7.48)
High School 1	0.00101***	-0.03427*	1.94596***	3.22851***	1.86782***
	(4.97)	(-1.97)	(8.10)	(4.26)	(8.58)
High School 2		0.03052***	8.93532*	20.21510*	8.07884*
		(4.65)	(2.31)	(2.26)	(2.41)
High School μ			-0.65544	-0.53079***	-0.64217***
			(-9.99)	(-12.42)	(-10.49)
Safety 1	0.00016**	0.02600	85.62664*	10.08956***	3.18420***
	(2.61)	(1.27)	(2.29)	(4.14)	(8.07)
Safety 2		0.00180	2.70518	0.81677	0.24090
		(0.43)	(0.78)	(0.84)	(0.85)
Safety μ			-0.28397***	-0.36102***	-0.40195***
			(-9.59)	(-7.54)	(-6.57)
Distance to Hazard 1	0.02576 ***	0.07352**	0.00640**	0.08105***	0.28460***
	(3.55)	(3.17)	(2.71)	(4.56)	(8.66)
Distance to Hazard 2		-0.11634*	96.01543	99.52364	82.27298
		(-2.24)	(1.49)	(1.36)	(1.31)
Distance to Hazard μ			-0.22796***	-0.27502***	-0.30074***
			(-8.36)	(-6.18)	(-5.38)
Non Black 1	1.70E-06	0.04019***	0.00018***	0.01362***	0.11700***
	(1.58)	(8.43)	(5.33)	(9.19)	(17.37)
Non Black 2		-0.00388***	210.26070*	246.86850	248.21080
		(-8.00)	(2.10)	(1.90)	(1.80)
Non Black μ			-0.25031***	-0.31726***	-0.35721***
			(-24.32)	(-16.72)	(-13.81)
Non Hispanic 1	2.26E-06*	0.03730***	0.00042***	0.02017***	0.14098***
	(2.45)	(6.74)	(7.51)	(10.42)	(17.40)
Non Hispanic 2		-0.00346***	151.11420*	153.42230*	141.94710*
		(-6.26)	(2.46)	(2.34)	(2.23)
Non Hispanic μ			-0.23872***	-0.30279***	-0.34135***
			(-25.62)	(-18.01)	(-14.68)

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

	Standard Model[a]	All μ's equal -1	Estimated μ's		
			σ_3 = 1/2[b]	σ_3 =1	σ_3 =2
Commute 1	24.4689**	30.14236***	0.37957**	0.37279**	0.37541**
	(2.78)	(3.41)	(3.11)	(2.97)	(3.02)
Commute 2	-25.07227**	-31.10109***	0.25742***	0.26160***	0.26042***
	(-2.64)	(-3.26)	(3.86)	(3.80)	(3.82)
Income Tax Rate	5.67241	0.06265	0.95973	0.30414	0.61753
	(1.45)	(0.02)	(0.22)	(0.07)	(0.14)
School Tax Rate	1.40680	-1.17535	1.91474	2.08131	1.97956
	(1.35)	(-1.01)	(1.59)	(1.70)	(1.63)
City Tax Rate	-1.48317	-2.55823*	3.35662*	3.54533*	3.51780*
	(-1.13)	(-1.98)	(2.38)	(2.49)	(2.48)
Tax Break Rate	2.42167	5.39294**	-6.76149***	-6.98650***	-6.95556***
	(1.30)	(2.93)	(-3.21)	(-3.28)	(-3.28)
Not City	0.01335	0.04731	0.06343*	0.06262*	0.06362*
	(0.43)	(1.51)	(2.20)	(2.17)	(2.20)
SSE	42.9032	38.9536	36.0778	36.1882	36.2337
R-squared	0.6354	0.6689	0.6934	0.6924	0.6920
Root MSE	0.1630	0.1556	0.1500	0.1503	0.1504

Notes: The dependent variable is the CBG fixed effect (in log form); the number of observations is 1665; the regressions in this table also include the variables in Table 6 plus worksite and county fixed effects. t-statistics are in parentheses; a * (**) [***] indicates significance at the 5 (1) [0.1] percent level.

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

[b] This regression sets σ_3 = 2 for High School.

Table 6: Results for Geographic Control Variables

	Standard Model[a]	All μ's equal -1	Estimated μ's σ_3 =1/2[b]	σ_3 =1	σ_3 =2
Lakefront	0.11499***	0.09339***	0.08307**	0.08313**	0.08301**
	(5.05)	(4.25)	(3.07)	(3.06)	(3.05)
Distance to Lake	-0.06555***	-0.04278**	-0.03420	-0.03353	-0.03395
	(-3.81)	(-2.57)	(-1.72)	(-1.68)	(-1.70)
Snowbelt 1	0.03559***	0.03183***	0.03986***	0.03968***	0.03931***
	(4.87)	(4.49)	(4.45)	(4.42)	(4.37)
Snowbelt 2	-0.00168***	-0.00133**	-0.00203***	-0.00202***	-0.00200***
	(-3.63)	(-2.95)	(-3.84)	(-3.80)	(-3.77)
Ghetto	-0.10671***	-0.09680***	-0.09191**	-0.09177**	-0.09150**
	(-5.59)	(-4.80)	(-2.65)	(-2.66)	(-2.66)
Near Ghetto	-0.06971**	-0.07622***	-0.08050***	-0.08389***	-0.08425***
	(-3.11)	(-3.43)	(-3.65)	(-3.78)	(-3.80)
Near Airport	0.08853*	0.08695*	0.06172	0.06583	0.06644
	(2.18)	(2.20)	(1.77)	(1.90)	(1.91)
Airport Distance	-0.01656***	-0.01831***	-0.01206**	-0.01231**	-0.01247**
	(-3.37)	(-3.86)	(-2.92)	(-2.97)	(-3.00)
Near Public	0.00830	0.00110	0.00310	0.00318	0.00340
	(0.70)	(0.10)	(0.31)	(0.31)	(0.34)
Near Private	0.02508*	0.01177	0.00997	0.01157	0.01181
	(2.05)	(0.10)	(0.77)	(0.89)	(0.91)
Land Release	-0.00001	0.00001	-4.09E-06*	-4.14E-06*	-0.00000*
	(-0.94)	(0.50)	(-1.99)	(-2.10)	(-2.23)
Smog NE	-0.24431**	-0.28107***	-0.39884***	-0.40433***	-0.39802***
	(-2.69)	(-3.20)	(-4.59)	(-4.64)	(-4.54)
Smog SE	-0.13161	--0.11739	-0.23566**	-0.24029**	-0.23545**
	(-1.57)	(-1.44)	(-2.74)	(-2.78)	(-2.71)
Smog SW	-0.23139**	-0.25828**	-0.31323***	-0.31845***	-0.31497***
	(-2.61)	(-3.04)	(-3.77)	(-3.80)	(-3.77)
Smog NW	0.04340	0.06348	-0.04950	-0.05911	-0.05644
	(0.22)	(0.33)	(-0.37)	(-0.42)	(-0.4)
Smog Direction NE	0.01191*	0.01317**	0.01983***	0.02007***	0.01974***
	(2.46)	(2.81)	(4.17)	(4.20)	(4.11)
Smog Direction SE	0.00371	-0.00069	0.01165*	0.01171*	0.01144*
	(0.63)	(-0.12)	(2.04)	(2.04)	(1.99)
Smog Direction SW	0.01032	0.01148*	0.01459**	0.01482**	0.01476**
	(1.88)	(2.18)	(2.93)	(2.95)	(2.94)
Smog Direction NW	-0.11092	-0.14053	-0.08356	-0.08072	-0.08042
	(-1.06)	(-1.40)	(-1.43)	(-1.34)	(-1.34)
Local Amenities	0.01965*	0.01854*	0.01624*	0.01642*	0.01665*
	(2.30)	(2.26)	(2.03)	(2.05)	(2.07)

Table 6: Results for Geographic Control Variables (continued)

	Standard Model[a]	All μ's equal -1	Estimated μ's σ_3 =1/2[b]	σ_3 =1	σ_3 =2
Freeway	0.00005	0.00449	0.00894	0.00819	0.00792
	(0.00)	(0.36)	(0.74)	(0.68)	(0.65)
Railroad	-0.04472***	-0.03160**	-0.02932**	-0.02893**	-0.02879**
	(-4.21)	(-3.04)	(-2.76)	(-2.71)	(-2.69)
Shopping	-0.00422	-0.00664	-0.00984	-0.00908	-0.00878
	(-0.41)	(-0.67)	(-1.04)	(-0.96)	(-0.93)
Hospital	-0.02186*	-0.01851	-0.01323	-0.01391	-0.01411
	(-2.10)	(-1.85)	(-1.29)	(-1.36)	(-1.38)
Small Airport	0.02467	0.02838	0.03641	0.03850	0.03919
	(0.89)	(1.07)	(1.69)	(1.78)	(1.81)
Big Park	0.03156*	0.01253	0.01248	0.01206	0.01245
	(2.38)	(0.97)	(1.19)	(1.15)	(1.19)
Historic District	0.00981	0.02281	0.01778	0.01892	0.01970
	(0.62)	(1.49)	(0.93)	(0.98)	(1.02)

Notes: The dependent variable is the CBG fixed effect (in log form); the number of observations is 1665; the regressions in this table also include the variables in Table 5 plus worksite and county fixed effects. t-statistics are in parentheses; a * (**) [***] indicates significance at the 5 (1) [0.1] percent level.

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

[b]This regression sets σ_3 = 2 for High School.

Table 7. Tests of the Normal Sorting Hypothesis

Explanatory Variable	Elementary Score	9th Grade Score	High School	Safety	Hazard	Non Black	Non Hispanic
Log of Family Income if Slope	-0.83636***	1.24066***	0.62270**	0.88655***	0.47830**	4.6081***	-0.39267
of Bid Function is Positive	(-3.47)	(8.06)	(3.01)	(3.71)	(2.61)	(3.91)	(-0.32)
Log of Family Income if Slope			-1.12175***	0.08069*		-0.05067	
of Bid Function is Negative			(-6.87)	(2.09)		(-0.36)	
Number of Observations	1665	1634	1665	1665	1658	1665	1650
Number of Uncensored Obs.	258						
Wald Test of Independent Equations							
$chi^2(1)$	6611.00	5.52		712.00		1.42	
Pr > chi^2	0.0000	0.0188		0.0076		0.2339	

Notes: These results are estimates of (11) with the slope of the bid function (ψ) as the dependent variable. The estimations are based on OLS (9^{th} Grade Score, Distance to Hazard, Non Hispanic), a selection model (Elementary Score), or an endogenous switching model (High School, Safety, Non Black). The associated selection equations are presented in Table 8. The explanatory variables other than income, which differ across columns, come from the list of variables in Table 8 but are not presented here. The OLS estimations each drop a few observations with negative values for ψ. t-statistics are in parentheses; a * (**) [***] indicates significance at the 5 (1) [0.1] percent level. Full results are available from the author upon request.

Table 8: Selection Equations

Block Group Trait	Elementary Score	9th Grade Score	Safety	Non Black
Median family income (log)	0.11896	-1.47600***	-1.28429***	-1.04272**
	(0.81)	(-4.72)	(-5.00)	(-2.57)
Percent married	0.00575*	-0.00720	-0.03063***	-0.05332***
	(2.08)	(-1.20)	(-5.57)	(-5.43)
	-0.03050***			
	(-4.30)			
Percent of households with kids		-0.01738**	-0.02508***	-0.00850
		(-2.68)	(-3.79)	(-0.83)
Percent of kids in private elementary	0.00584***	0.01820***		-0.02642***
	(4.54)	(6.74)		(-4.09)
Percent of kids in private high school		0.01174***		0.00125
		(4.62)		(0.24)
Percent with English as 2nd language	-0.01093***	-0.02552*	0.00592	0.20707***
	(-3.34)	(-2.55)	(0.84)	(6.16)
Percent foreign born		-0.05172**	0.01180	-0.04678
		(-3.10)	(0.85)	(-0.93)
Percent Asian and Pacific Islander		-0.03725	0.04802*	0.19734***
		(-1.75)	(2.06)	(3.75)
Percent Catholic (county)	-0.01324	0.05137***		
	(-1.92)	(5.29)		
Percent other Christian denomination	-0.02889**	0.04788**		
	(-2.66)	(2.99)		
Index of educational homogeneity	-2.13466**			
	(-2.98)			
Percent in house < 1 year (tract)	0.00383	0.01977*	0.00718	-0.05079*
	(0.53)	(2.16)	(0.75)	(-2.54)
Percent owner occupied		0.00340	0.01991***	0.01514*
		(0.95)	(5.58)	(2.46)

Table 8: Selection Equations (continued)

Block Group Trait	Elementary Score	9th Grade Score	Safety	Non Black
Percent blue collar		0.00542	-0.00098	-0.06070***
		(0.80)	(-0.17)	(-5.42)
Unemployment rate		0.05416***	0.04455***	0.08390***
		(3.83)	(3.67)	(4.53)
Percent with high school degree only	0.00934*	-0.01375	-0.03792***	-0.03164*
	(2.27)	(-1.47)	(-4.69)	(-2.14)
Percent with some college	0.00960*	-0.02084*	-0.04420***	0.02516
	(2.27)	(-2.09)	(-4.85)	(1.65)
Percent with college degree	0.00355*	-0.03672***	-0.01483	-0.09812***
	(0.71)	(-3.33)	(-1.47)	(-4.71)
Percent with graduate degree	0.00624	-0.06575***	-0.04649***	-0.00380
	(1.16)	(-5.10)	(-3.58)	(-0.20)
Percent elderly		-0.01904*	-0.04145***	-0.01291
		(-2.37)	(-5.19)	(-1.03)
Percent minority (school district)	-0.00391**			
	(-2.72)			
Percent on welfare (school district)	0.02488***			
	(9.39)			
Percent of housing in rural area	-0.00458**			
	(-3.09)			
	-0.25550			
	(-0.15)			

Notes: These results describe the selection equations associated with demand estimates in Table 7. The dependent variable is whether the bid function slope, ψ, is positive or negative. For Elementary Score the coefficients indicate selection into a positive slope; for the other columns, the coefficients indicate selection into a negative slope. The explanatory variables apply to the CBG unless another unit of geography is indicated. t-statistics are in parentheses; a * (**) [***] indicates significance at the 5 (1) [0.1] percent level. Full estimation results are available from the author upon request.