TECHNICAL WORKING PAPER SERIES

ESTIMATING HEDONIC MODELS: IMPLICATIONS OF THE THEORY

Helen Tauchen
Ann Dryden Witte

Technical Working Paper 271
http://www.nber.org/papers/T0271

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2001

Estimating Hedonic Models: Implications of the Theory
Helen Tauchen and Ann Dryden Witte
NBER Technical Working Paper No. 271
July 2001
JEL No. D4, H0, I0, J3

ABSTRACT

In this paper we consider the conditions under which instrumental variables methods are required in estimating a hedonic price function and its accompanying demand and supply relations. We assume simple functional forms that permit an explicit solution for the equilibrium hedonic price function. The principles are the same for models in which no analytic solution exists, but having the solutions makes the issues far more transparent. The need for instrumental variables estimation is directly analogous for the classical demand and supply model with undifferentiated products and for the hedonic model with differentiated products. In estimating individual demand and supply functions, instrumental variables estimation is required if the consumer and firm unobservables, which give rise to the error terms in the demand and supply functions, are correlated across consumers/firms within a community. In estimating inverse demand/supply functions, which are referred to as bid/offer functions in the hedonic model, instrumental variables estimation is required even if the unobservables are not correlated across agents within a community. If the unobservables are not correlated across agents within a community, then community binaries or the means of observable consumer and firm characteristics can be used as instruments. If the unobservables are correlated then only the latter can be used.

The error term in the hedonic price function is often assumed to be uncorrelated with the chosen attributes. This assumption may be reasonable if consumers have quasilinear preferences. If not, then the error term in the price function may affect the utility-maximizing amounts of the attributes. The feasible instruments again depend upon whether the error term is correlated for agents within a community. If not, then community binaries or observed individual characteristics may be used as instruments. If so, then the community binaries are correlated with the error terms and cannot serve as instruments.

Helen Tauchen
Department of Economics
University of North Carolina
tauchen@unc.edu

Ann Dryden Witte
Department of Economics
Wellesley College
and NBER
awitte@wellesley.edu

1. Introduction

In this paper, we consider a simple hedonic model with a differentiated product. Our assumed utility and cost functions yield an explicit expression for the equilibrium hedonic price function. Based on this theoretical model, we derive an empirical specification with both observable and unobservable consumer and firm characteristics. The errors in the hedonic demand/supply and bid/offer functions are attributable to the unobservable characteristics. With this approach, we can show very clearly how the endogeneity issues that arise in estimation of hedonic models are related to those of the standard model for an undifferentiated good. The principles are identical for cases in which the equilibrium price function cannot be derived analytically, but having explicit solutions makes the relationships far more transparent.

The focus of this paper is on the endogeneity issues that arise in the estimation of hedonic demand and supply functions and of the equilibrium hedonic price function. We are particularly interested in empirical applications using data from more than one community. The general literature on hedonics is vast and covers many other aspects of the hedonic model including parametric versus nonparametric estimation (Anglin and Gencay, 1996), heteroscedasticity of the error terms (Yoo, 2001), principal components analysis for dealing with collinearity of the attributes, and functional form choice (Berndt and Showalter, 1993), none of. which we address. Cheshire and Sheppard (1998) provide an excellent review; they are interested specifically in the housing market but the same issues are relevant in other contexts.

Most of the work on endogeneity and hedonic models deals with the estimation of the demand/supply or the bid/offer functions rather than of the hedonic price function itself. Some of the initial research on empirical applications included suggestions for using trading partner characteristics for estimating the hedonic demand and supply relations. The reasons that these instruments are unsatisfactory have been well-explained (Kahn and Lange, 1988). Instruments now considered include community binaries and community characteristics such as climate and road systems that affect the desirability of a location and hence the equilibrium price function but do not themselves enter the demand or supply relations. Deriving the empirical specification directly from the theoretical model makes it straightforward to determine the conditions under which these variables and others are appropriate instruments.

Far less attention has been focused on whether or not the estimation of the hedonic price function involves any endegeneity concerns. Wooldridge (1996) suggests using the individual consumer and firm characteristics (e.g., income, education, input prices) as instruments in estimating the hedonic price function but many applications do not mention the issue. Again, having an empirical specification that follows directly from the theoretical model makes it straightforward to explain the conditions under which the chosen product attributes are correlated with the error term in the hedonic price function. Not unsurprisingly, the possible instruments depend upon whether or not the unobservables are correlaed for consumers and firms within the same community.

As explained later in the paper, there is a direct analogy between the estimation of demand and supply functions for undifferentiated products and of hedonic demand and supply functions. In the standard model, the quantities supplied and demanded depend upon the observed market prices; in the hedonic model the attributes supplied and demanded depend upon the parameters of the hedonic price function. The conditions under which endogeneity arises are precisely the same for the standard and the hedonic models, namely, whenever the consumer/firm unobservables, which give rise to the error terms in the individual demand/supply functions, are correlated. With data from different communities, the natural instruments for estimating the standard and the hedonic demand and supply functions include community characteristics such as average income or input prices. These factors may affect the equilibrium price or price function but do not enter the individual consumer or firm demand functions directly.

For practical reasons related to the numbers of parameters of the price functions, the hedonic demand/supply functions are seldom estimated. A more common approach is to estimate bid (offer) functions with the marginal prices paid by a consumer (offered by a firm) dependent upon the chosen product attributes and the consumer (firm) characteristics. This is analogous to estimating standard inverse demand (supply) functions with the prices dependent upon the quantities demanded (supplied) and individual consumer (firm) characteristics. There are two potential sources of endogeneity in estimating both inverse demand/supply and bid/offer and functions for individual agents. The first is the same as for direct supply and demand models of undifferentiated products and occurs whenever there are unobservable individual characteristics that are correlated across agents within a community. The second is peculiar to the inverse demand/supply and bid/offer models in which the chosen attributes are right-hand side variables in the empirical model. With this set-up, the error terms in the individual indirect demand/supply and bid/offer functions are necessarily correlated with the chosen product attributes. The appropriate instruments depend upon whether both sources of endogeneity or only the second are present. Having a model with an explicit solution makes it very easy to understand in which of these cases the commonly-suggested instruments (e.g., community binaries) are valid.

With the hedonic model, another possibility is the estimation of modified demand (supply) functions. The basic approach is to model the attributes demanded (supplied) as dependent upon the estimated marginal prices and individual consumer (firm) attributes. Although the form of the functions are quite different, the endogeneity issues and the feasible instruments are the same for the modified demand (supply) and the bid (offer) functions.

A fundamental difference between the hedonic and standard models is that the price in the standard model is observed whereas the hedonic price function must be estimated. The key endogeneity issue for estimation of the hedonic price function is not directly related to those for the demand/supply or bid/offer functions. For the hedonic price function, the issue is whether the price function error term represents factors that are unobserved by both the agents and the researcher or only by the researcher. In the latter case, the error term may be correlated with the product attributes, and instrumental variables estimation is required.

The organization of the remainder of the paper is as follows. In the next section, we describe a hedonic model in which we assume explicit functional forms for the utility and cost functions and for the distribution of consumer and firm characteristics. Given these functional forms, we derive the equilibrium hedonic price function. In the third section, we specify the empirical model in which consumers and firms have unobservable characteristics that give rise to the error terms in their demand/supply functions and in the hedonic price function. In this section, we also describe the endogeneity issues involved in estimating the individual demand/supply or bid/offer functions. In the following section, we describe the same issues for the hedonic price function. In all cases, the appropriate instruments depend upon the correlation structures of the unobservables. The paper concludes with a short summary.

2. Hedonic Model

We use the standard hedonic model with consumers and producers of a differentiated product and derive the equilibrium price function for the product. The functional forms for the consumers' utility functions and the firms' cost functions are chosen to yield analytic solutions for the price function. The arguments of the demand and supply functions and of the equilibrium hedonic price function are the same for utility and cost functions that do not allow an explicit expression for the equilibrium hedonic price function.

2.1 Consumers

The model is very similar to that of Epple (1987) and others. Each individual consumes one unit of commodity Z with attributes $z \epsilon R^N$. An individual's well-being depends on the attributes consumed and on expenditure on all other goods y. The utility function is quasilinear, and

$$U(z, y) = -\frac{(z-\alpha)' A (z-\alpha)}{2} + y$$

where $A \epsilon R^{NxN}$ and $\alpha \epsilon R^N$ represent the consumer taste and preference characteristics. In empirical work, these might depend on factors such as age or education. The matrix A is assumed to be symmetric and positive definite and is the same for all consumers. The vector α is distributed normally across the population with mean $\bar{\alpha}$ and variance V_α.

Let I denote income and $P(z)$ the price of a unit of the with attribute vector z. The utility of an individual with income I and attribute vector z is

$$-\frac{(z-\alpha)' A (z-\alpha)}{2} + I - P(z).$$

Each individual selects the attribute vector z to maximize utility. For an individual of type α, the first order conditions for an interior optimum are

$$\nabla P(z) = -A(z - \alpha). \tag{1}$$

Given the quasilinear utility function, these N functions are the usual bid equations.

2.2 Producers

Each firm produces at most one unit of good Z and has quadratic cost

$$C(z) = \beta z + \frac{z' B z}{2}$$

where $B \epsilon R^{NxN}$ is symmetric and positive definite and $\beta \epsilon R^N$. The matrix B and vector β are the parameters of the cost function and measure characteristics such as input prices and technology. All firms are assumed to have the same matrix B; the vector β is distributed normally with mean $\bar{\beta}$ and variance $V_{\beta.}$ The assumptions that A and B do not vary across agents is necessary in order to obtain an explicit solution for the equilibrium hedonic price function but are not required for the general hedonic model.

A firm's profit for attribute vector z is $P(z) - C(z)$. For a firm of type β, the first order conditions for an interior optimum are

$$\nabla P(z) = \beta + Bz, \tag{2}$$

and these are the N offer equations.

2.3 Equilibrium Hedonic Price Function

In this section we show that a quadratic price function of the form

$$P(z) \equiv \gamma' \ z + \frac{z' \ \Gamma \ z}{2}$$

satisfies the definition of an equilibrium, and we determine the how $\gamma \epsilon R^N$ and $\Gamma \epsilon R^{NxN}$ are related to the parameters of the utility and cost functions. For the quadratic price function, we can solve the bid and offer equations explicitly for the amounts of the attributes demanded and supplied. The bid equation (1) is $\gamma + \Gamma z = -A(z - \alpha)$. Solving for the z vector demanded yields

$$z^d = (\Gamma + A)^{-1}[A\alpha - \gamma]. \tag{3}$$

The second order part of the sufficient condition for an interior optimum to the consumer's problem is that the matrix $-(\Gamma + A)$ is negative definite. Similarly, with the quadratic price function, the offer equation (2) is $\gamma + \Gamma z = \beta + Bz$. Solving for the z vector supplied yields

$$z^s = (\Gamma - B)^{-1}[B\beta - \gamma]. \tag{4}$$

The second order part of the sufficient condition for an interior optimum is that the matrix (Γ - B) is negative definite.

In this model, the parameters A, B, Γ, and γ appearing in the demand and supply expressions are the same for all consumers and firms. The differences in the attribute choices across consumers and firms is attributable to the differences in α and β which are assumed to be distributed normally. Since the characteristics α and β appear linearly in (3) and (4), the distribution of the z vectors demanded and supplied are also normal. In this case, a necessary and sufficient condition for equilibrium is that the means of the demand and supply vectors are equal, or,

$$(\Gamma + A)^{-1}[A\bar{\alpha} - \gamma] = (\Gamma - B)^{-1}[B\bar{\beta} - \gamma],$$

and that the respective variances are equal, or

$$(\Gamma + A)^{-1} A V_\alpha A (\Gamma + A)^{-1} = (\Gamma - B)^{-1} B V_\beta B (\Gamma - B)^{-1}$$

where the symmetry of V_α, V_β, A, B and Γ is exploited in avoiding the transpose notation.

Consider first the condition for equality of the variances which can be written as

$$(\Gamma + A)^{-1} V_A (\Gamma + A)^{-1} = (\Gamma - B)^{-1} V_B (\Gamma - B)^{-1} \tag{5}$$

where $V_A \equiv A V_\alpha A$ and $V_B \equiv B V_\beta B$. If V_α, V_β, A and B are all diagonal, then the solution is very simple. In this case, the matrices V_A and V_B are also diagonal and positive definite. Define $V_A^{1/2}$ and $V_B^{1/2}$ as the diagonal matrices consisting of the positive square roots of each of the corresponding terms in V_A and V_B. As noted by Epple (1987), the matrix

$$\Gamma = (I + V_B^{1/2} V_A^{-1/2})^{-1} (B - V_B^{1/2} V_A^{-1/2} A)$$

solves (5). The parameters of the squared terms in the equilibrium hedonic price function depend on the variances of household and firm characteristics and on the parameters A and B of the utility and cost functions. It is easily shown that the matrices -$(\Gamma + A)$ and $(\Gamma - B)$ are necessarily negative definite which is consistent with the second-order part of the conditions for the consumer and firm optimizations. There may be other solutions to (5) but they do not necessarily satisfy the definiteness conditions. Given the assumed distribution of consumer and firm characteristics, all firms and consumers participate in the market.[1]

If V_α, V_β, A and B are not all diagonal, then the solution is less straightforward but the principle is the same. The matrices V_A and V_B are both positive definite since V_α and V_β are variances. Both V_A and V_B therefore have nonsingular square roots but the square roots are not necessarily unique or symmetric. Let S_A and S_B denote square root matrices of A and B; by definition, $S_A S_A' = V_A$ and $S_B S_B' = V_B$. The candidates for the solution to (5) are of the form

$$\Gamma_c = (I - S_B S_A^{-1})^{-1} (B + S_B S_A^{-1} A). \tag{6}$$

In order for one of the candidates Γ_c to be a solution it must be symmetric and the matrices $(I - S_B S_A^{-1})^{-1}(A + B)$ and -$(I - S_B S_A^{-1})^{-1}$ $S_B S_A^{-1}(A + B)$ must be positive definite. Although solutions do not necessarily exist for all parameter values, there are open subsets of parameters for which they do.[2]

Finally, the equality of the means of the z vectors demanded and supplied imply that

$$\gamma = [(\Gamma - B)^{-1} - (\Gamma + A)^{-1}]^{-1} [(\Gamma - B)^{-1} B \bar{\beta} - (\Gamma + A)^{-1} A \bar{\alpha}]. \tag{7}$$

The parameters of the linear terms in the hedonic price function depend on the means and variances of the household and firm characteristics in addition to their variances. Although the exact nature of the relationship between the parameters of the utility and cost functions and the parameters of the hedonic price function are particular to the assumed functional forms of the utility and cost functions, the general principle holds for any utility and cost functions. Namely, the hedonic price functions differs across markets in accordance with the distribution of the characteristics of the demanders and suppliers. In this case, only the means and variances of the distributions affect the parameters of the price function; for other distributions the parameters of the equilibrium price function might also depend on higher moments.

3 . Estimation of Demand/Supply Relations

To incorporate the error terms in the empirical model, we assume that the characteristic vector α in the utility function has three components. Specifically, the vector α_{ij} of characteristics for consumer i in community j is

[1] Alternatively, the distribution function for consumers might be N_c times the assumed normal distribution function for α where N_c is a measure of the number of consumers. Similarly, the distribution function for firms might be N_f times the assumed normal distribution function for β. If $N_c \neq N_f$, then some agents do not participate in the market and the solution involves boundary conditions describing which agents participate and which do not.

[2] For example, if $\bar{\alpha} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\bar{\beta} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $V_\alpha = \begin{bmatrix} 3 & -.3 \\ -.3 & 2 \end{bmatrix}$, $V_\beta = \begin{bmatrix} 3 & -.1 \\ -.1 & 1 \end{bmatrix}$, $A = \begin{bmatrix} 3 & .1 \\ .1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 4 & .2 \\ .2 & 5 \end{bmatrix}$, then $\gamma = \begin{bmatrix} 3.183 \\ 2.046 \end{bmatrix}$ and $\Gamma = \begin{bmatrix} 2 & .1 \\ .1 & 1 \end{bmatrix}$. Small changes in any of the exogenous variables cause the parameters of the equilibrium price function to change in a continuous fashion.

$$\alpha_{ij} \equiv \alpha^o_{ij} + \mu^\alpha_{ij} + \varepsilon^\alpha_{ij}$$

where α^o_{ij}, μ^α_j, and ε^α_{ij} are each vectors in R^n. The first component (α^o_{ij}) denotes consumer characteristics that are observable to the researcher, the second (μ^α_{ij}) denotes characteristics that are unobservable and are correlated across consumers within community j, and the third $(\varepsilon^\alpha_{ij})$ denotes characteristics that are unobservable and uncorrelated across consumers. In the simplest case, the component μ^α_{ij} might be the same for all consumers in a community or $\mu^\alpha_{ij} = \mu^\alpha_j$. For some characteristics, only the observable component might be nonzero. For others, only the unobservables or some combination of the observable and the two unobservables might be nonzero. Some of the observables could be the same for all consumers in the community. For direct comparison with the theoretical model, we assume that the observable and unobservable characteristics are distributed normally across any community but are not necessarily the same for all communities. We also assume that the observable consumer characteristics are uncorrelated with the unobservable consumer characteristics. Similarly, the characteristic vector $\beta_{kj} \epsilon R^n$ for firm k in community j has three components $\beta_{kj} \equiv \beta^o_{kj} + \mu^\beta_{kj} + \varepsilon^\beta_{kj}$ with the same interpretation as for consumers. The means of the consumer and producer characteristics for community j are denoted $\bar{\alpha}_j$ and $\bar{\beta}_j$ respectively and the variances are denoted $V_{\alpha j}$ and $V_{\beta j}$.

As derived in the previous section, the parameter vector of the linear term in the equilibrium hedonic price function for community j is

$$\gamma_j = M^\alpha_j \ \bar{\alpha}_j + M^\beta_j \ \bar{\beta}_j = M^\alpha_j \ (\bar{\alpha}^o_j + \bar{\mu}^\alpha_j + \bar{\varepsilon}^\alpha_j) + M^\beta_j \ (\bar{\beta}^o_j + \bar{\mu}^\beta_j + \bar{\varepsilon}^\beta_j) \tag{8}$$

where

$$M^\alpha_j = -(\Gamma_j - B)^{-1} - (\Gamma_j + A)^{-1}]^{-1}(\Gamma_j + A)^{-1}A \tag{9}$$

and

$$M^\beta_j = (\Gamma_j - B)^{-1} - (\Gamma_j + A)^{-1}]^{-1}(\Gamma_j - B)^{-1}B \tag{10}$$

If the matrices A_j and B_j are diagonal and each of the consumer and firm characteristics is distributed independently, then the parameter matrix for the quadratic term of the hedonic price function is

$$\Gamma_j = (I + V_B^{1/2} V_A^{-1/2})^{-1}(B - V_B^{1/2} V_A^{-1/2} A) \tag{11}$$

where $V_{Aj} \equiv A V_{\alpha_j} A$ and $V_{B_j} \equiv B V_{\beta_j} B$. If the matrices A, B, V_{Aj}, and V_{Bj} are not all diagonal, then there is no closed form solution but Γ_j depends on these same four matrices. Assuming no government regulations or unusual geographic features, the hedonic price function for community j is then

$$P^j(z) = \gamma'_j \ z + \frac{z' \ \Gamma_j \ z}{2}. \tag{12}$$

A common approach for estimating demand/supply or bid/offer functions in the hedonic literature is a two-stage procedure with the first stage being the estimation of the hedonic price function. In this section, we adopt the two-step approach and focus on the estimation of the hedonic demand and supply functions assuming that the parameters of the hedonic price function have been estimated. In subsequent sections, we discuss the estimation of the hedonic price function together with the demand and supply functions.

3.1 Demand & Supply Functions

Written explicitly in terms of the components of α_{ij} and β_{kj}, the demand by individual i in community j as given by (3) is

$$z^d_{ij} = (\Gamma_j + A)^{-1}[A(\alpha^o_{ij} + \mu^\alpha_{ij} + \varepsilon^\alpha_{ij}) - \gamma_j], \tag{13}$$

and the supply by firm k in community j as given by (4)

$$z^s_{kj} = (\Gamma_j + B)^{-1}[B(\beta^o_{kj} + \mu^\beta_{kj} + \varepsilon^\beta_{kj}) - \gamma_j]. \tag{14}$$

Note that these functions are exactly analogous to the usual demand and supply functions except that the parameters of the hedonic price function, γ_j and Γ_j, rather than the prices themselves enter. The error terms of the demand and supply functions are attributable to the unobservables μ_{ij}^{α}, $\varepsilon_{ij}^{\alpha}$, μ_{kj}^{β} and ε_{kj}^{β} . The parameters to be estimated are the elements of the symmetric matrices A and B.

The source of the endogeneity in estimating the demand equation is the correlation of the error term μ_{ij}^{α} with the cost function parameter γ_j appearing as an argument in the demand function. Since μ_{ij}^{α} is correlated across consumers, μ_{ij}^{α} is correlated nontrivially with $\bar{\mu}_j^{\alpha}$ which is one of the determinants of γ_j. Consider, for example, the error components model in which μ_{ij}^{α} is the same for each consumer or $\mu_{ij}^{\alpha} = \bar{\mu}_j^{\alpha}$. The term μ_{ij}^{α} which appears in the demand function is obviously correlated with the term $\bar{\mu}_j^{\alpha}$ which appears in γ_j as given by (8). The second part of the error term does not generate endogeneity complications. In a community with many consumers, the independently distributed portion of the error term in the demand equation ($\varepsilon_{ij}^{\alpha}$) would have only a negligible correlation with $\bar{\varepsilon}_j^{\alpha}$ and hence with the parameter vector γ_j appearing in the demand equation. For the hedonic supply function, the endogeneity arises in an exactly analogous fashion from the correlation between the error term μ_{kj}^{β} and the price function parameter γ_j appearing in the supply function.

The nature of the endogeneity problems in the estimation of simple demand and supply functions for homogeneous goods is exactly analogous to that for the hedonic model. In order to see clearly the parallel between the hedonic and the homogeneous good cases, it is helpful to consider in detail the usual supply and demand model. Assuming a linear functional form, the quantity of good X demand by individual i in community j is

$$x_{ij}^{d} = \mathrm{a}_0 + \mathrm{a}_p p_j + \mathrm{a}_\alpha(\alpha_{ij}^{o} + \mu_{ij}^{\alpha} + \varepsilon_{ij}^{\alpha}) \tag{15}$$

where p_j is the price of the good in community j and where a_0, a_p and a_α are the parameters of the demand function. The individual characteristics α_{ij}^{o}, μ_{ij}^{α}, and $\varepsilon_{ij}^{\alpha}$ have the same interpretation as for the hedonic model. This is a standard linear demand model with the error terms that may be correlated across individuals in the same community. Similarly, the quantity of good X supplied by firm k in community j is

$$x_{kj}^{s} = b_0 + b_p p_j + b_\beta(\beta_{kj}^{o} + \mu_{kj}^{\beta} + \varepsilon_{kj}^{\beta}) \tag{16}$$

where b_0, b_p and b_β are the parameters of the supply function. Solving for the equilibrium price in community j yields

$$p_j^{*} = \frac{N_j^{D}(a_0 + a_\alpha(\bar{\alpha}_j^{o} + \bar{\mu}_j^{\alpha} + \bar{\varepsilon}_j^{\alpha})) - N_j^{S}(b_0 + b_\beta(\bar{\beta}_j^{o} + \bar{\mu}_j^{\beta} + \bar{\varepsilon}_j^{\beta}))}{b_p N_j^{S} + a_\alpha N_j^{D}}$$

where N_j^{S} and N_j^{D} are the number of suppliers and demanders. The source of the endogeneity in estimating the demand equation is the same as for the hedonic model, namely the correlation between μ_{ij}^{α} which appears in the individual demand equations and $\bar{\mu}_j^{\alpha}$ which affects the equilibrium price.

The usual instruments for estimation of the standard demand equation are the means of the observable producer characteristics $\bar{\beta}_j^{o}$ which affect the equilibrium price but are not correlated with the error terms of the demand function and do not appear as arguments of the demand function. Similarly, the instruments for estimation of the supply equation are the means of the observable consumer characteristics $\bar{\alpha}_j^{o}$.

The instrumental variables approach for estimating the hedonic demand function is very similar to that for the homogenous-good supply and demand model. In order to estimate the hedonic demand function by instrumental variables, there must be factors that (i) do not appear in the demand equation and are not correlated with the error terms μ_{ij}^{α} and $\varepsilon_{ij}^{\alpha}$ in the demand equation but (ii) affect the parameter γ_j that is correlated with the error terms. The possible instruments are precisely the same as for the usual supply and demand model. In estimating the demand equation, the instruments are the means $\bar{\beta}_j^{o}$ and the variances $V_{\beta j}$ of the observable firm characteristics (Kahn and Lang, 1988). These variables affect γ_j but do not appear in the demand equation. Using the same arguments as in the classical demand and supply estimation, it may be reasonable to assume that the means and variances of the observable firm characteristics are not correlated with the error terms in the individual demand equations. Note that the appropriate instruments involve the distribution of the characteristics of the firms in the community rather than the characteristics of the particular firm from whom a consumer purchases the product. Binary variables for the communities are not, however, good candidates for instruments. Given the correlation across consumers in μ_{ij}^{α}, the community binaries would be correlated with the error terms in the demand equations.

An instrumental variables method might not be required if the only source of error in the demand and supply functions were the purely idiosyncratic errors $\varepsilon_{ij}^{\alpha}$ and ε_{kj}^{β}, or in other words if μ_{ij}^{α} and μ_{kj}^{β} were identically zero. In this case, the correlation between γ_j and the error terms in the supply/demand functions is due only to the correlation between $\varepsilon_{ij}^{\alpha}$ and its mean and similarly between ε_{kj}^{β} and its mean. Since the ε's are independently distributed, the correlation would be small in a community with many consumers and firms and might reasonably be ignored. In this case, instrumental variables methods are not required for either estimation of either the hedonic demand/supply functions or the simple supply and demand functions for a homogeneous good.

All of the standard questions related to identification and instrumental variables methods are relevant for the estimation of the hedonic demand /supply functions. Unless identification is based on functional form, one of the necessary conditions is that the number of variables excluded from the demand/supply or the bid/offer equations be at least as large as the number of included right-hand side variables for which there is correlation with the error terms. The quadratic form of the hedonic price function and the exact nature of the relationship between the parameters of the price function and the means and variances of the consumer and firm characteristics depends, of course, upon the assumed functional forms of the consumer's utility and firm's cost functions. For other functional forms, the equilibrium hedonic price function is not necessarily quadratic and may not even have a closed form solution. However, even in the more general case, the endogeneity issues and the appropriate choice of instruments remain the same.

3.2 Bid & Offer Functions

Direct estimation of the hedonic demand and supply functions poses a practical difficulty in models with many commodity attributes. Specifically, there are a large number of parameters of the hedonic price function and each of these coefficients appears as an argument in the demand and supply functions. In addition, the hedonic price function parameters are themselves estimated which complicates the computation of the standard errors of the estimated coefficients for the demand and supply functions. A more common approach in the hedonic literature is to estimate the bid and offer functions. There are two sources of endogeneity in estimating these functions. The first is the same as for estimation of the hedonic demand and supply functions, and the second is unique to the bid/offer function approach.

For the assumed utility and cost functional forms, the bid function (1) for consumer i in community j is

$$\gamma_j + \Gamma_j z_{ij}^d = -A(z_{ij}^d - \alpha_{ij}^o - \mu_{ij}^{\alpha} - \varepsilon_{ij}^{\alpha}) \ , \tag{17}$$

and the offer function (2) for firm k in community j is

$$\gamma_j + \Gamma_j z_{kj}^s = -B(z_{kj}^s - \beta_{kj}^o - \mu_{kj}^{\beta} - \varepsilon_{kj}^{\beta}). \tag{18}$$

The bid and offer functions are each a system of N equations. The variables z_{ij}^d and α_{ij}^o appearing on the right-hand side are observed; the left-hand side marginal prices are based on the previously estimated parameters of the hedonic price function. The endogeneity issues are attributable to the correlation between z_{ij}^d and the error terms μ_{ij}^{α} and $\varepsilon_{ij}^{\alpha}$. The first source of the correlation is the same as for the estimation of the demand function. Given the distributional assumptions for the error terms, μ_{ij}^{α} which appears in the demand function is correlated with $\bar{\mu}_j^{\alpha}$. The included right-hand side variable z_{ij}^d depends on γ_j (see 13) which depends on $\bar{\mu}_j^{\alpha}$ (see 8). The components of the matrices A and B are the parameters to be estimated.[3]

The second source of the correlation is that z_{ij}^d depends directly on μ_{ij}^{α} and $\varepsilon_{ij}^{\alpha}$ which appear as error terms in the bid equation. In contrast to hedonic demand function estimation, the endogeneity issue arises even if there is no correlation in the error terms across consumers in the same community. The correlation between the attributes chosen by a consumer and the idiosyncratic error term $\varepsilon_{ij}^{\alpha}$ still remains.

As for the estimation of the demand function, the means of the observed firm characteristics can serve as instruments for estimating the bid functions. The observed firm characteristics do not appear as arguments of the bid functions and are not correlated with the error terms of the bid functions. The means of the observed firm characteristics do, however, affect the parameters of the hedonic price function on which the optimal amount of the attributes depend, z_{ij}^d.

[3]If A and B differed across communities, then the parameters of the bid and offer functions would differ across communities.

As mentioned above, the endogeneity issue arises in estimating the bid function even if there is no correlation in the error terms across consumers. There are, however, additional possibilities for instruments. In this case, community binaries could be used. The community binaries would not be correlated with the error terms in the bid equations but would be correlated with the means of the firm characteristics. These characteristics affect the parameters of the hedonic price functions and hence the amounts of attributes chosen by consumers. Obviously, the use of the binaries rests heavily on the assumption that there are no unobserved community characteristics related to the consumers preferences for the goods (i.e., $\mu_{ij}^{\alpha} = 0$).

Estimating the bid and offer functions is analogous to estimating inverse demand and supply functions in the usual model with homogeneous goods. Let the inverse demand function for individual i purchasing the one undifferentiated good in community j be

$$p_{ij} = g_0 + g_x x_{ij}^d + g_\alpha \alpha_{ij}^o + \eta_{ij}^{\alpha}.$$

Any unobservable consumer characteristics that affect the demand for the good will be incorporated into the error term. Since the quantity demanded by the consumer depends on both the observable and unobservable characteristics, the error term is correlated with the quantity demanded by the consumer. This causes precisely the same difficulty as for the estimation of the bid function. In addition, the issues related to the correlation in the error term across individuals in a community are the same.

3.3 Modified Demand & Supply Functions

Another alternative to estimating the hedonic demand and supply functions is estimating modified demand and supply functions in which marginal prices of the attributes rather than the parameters of the hedonic price function appear as right-hand side variables.[4], Murray(1983) refers to such demand functions as mystical demand functions. The marginal prices for consumer i in community j are $\nabla P^j(z_{ij}^{d*})$ where z_{ij}^{d*} is the optimal consumption by the individual given the parameters of the hedonic price function. With the quadratic hedonic price function, $\nabla P^j\left(z_{ij}^{d*}\right) = \gamma_j + \Gamma_j z_{ij}^{d*}$. The left-hand side variables in the estimation of the modified demand functions are the same as for the hedonic demand itself, namely the attributes chosen by the consumers. The right-hand side variables are the observed consumer characteristics and the marginal prices.[5] Given the assumed functional forms, the modified demand functions are

$$z_{ij}^d = \alpha_{ij}^o + \mu_{ij}^{\alpha} + \varepsilon_{ij}^{\alpha} - A^{-1} P_{ij}^m \tag{19}$$

where the marginal price $P_{ij}^m \equiv P^j(z_{ij}^{d*})$. The endogeneity issues are exactly the same as for estimation of the bid functions. The error terms of the modified demand functions involve μ_{ij}^{α} and $\varepsilon_{ij}^{\alpha}$. These errors are correlated with the marginal prices since the prices depend on the chosen attribute vector z_{ij}^d which in turn depends on the error terms (see 13).

Evaluating the marginal price functions at the mean characteristics for community j presents additional issues concerning the relationship between the modified and true demand functions but does not necessarily avoid the endogeneity problems. The error terms μ_{ij}^{α} and μ_{kj}^{β} in the individual demand and supply equations are correlated with the means $\overline{\mu}_j^{\alpha}$ and $\overline{\mu}_j^{\beta}$. The parameters of the hedonic price function depend on these means and the attributes chosen depend upon the price function parameters. The endogeneity problem is the same as for the bid/offer functions. Only if there is no correlation in the error terms across individuals in a community, is the endogeneity problem avoided. In this case, the only source of correlation between the error terms in the demand equation and the marginal prices is attributable to the correlation between $\varepsilon_{ij}^{\alpha}$ and its mean. In a large community, this correlation would be negligible. Just as was the case for the unmodified demand equation, instrumental variables estimation is not required.

In recent work, Cheshire and Sheppard (1998) suggest using the marginal attribute prices paid by similar consumers as instruments in estimating the modified demand functions. In their work on housing demand functions, they based the definition of "similar" on both observed characteristics of consumers and geographic location. This approach is particularly attractive because it could be potentially used with data from only

[4]With this approach, the possibly nonlinear budget constraint is transformed to a linear budget constraint. The conditions under which this approach is valid have been well-investigated in the literature (e.g., Moffitt, 1989).

[5] If the utility function is not quasilinear and the attribute demand depends upon income, then the arguments of the modified demand functions also include an adjusted income term.

one community. Investigating the conditions under which instruments can be based on information from similar consumers is worthwhile.

First consider basing the definition of similarity on the observed characteristics of consumers in the same community. Substituting the attribute demands (3) yields the marginal prices

$$P_{ij}^{m} = \gamma_j + \Gamma_j\{(\Gamma_j + A)^{-1}[A(\alpha_{ij}^{o} + \mu_{ij}^{\alpha} + \varepsilon_{ij}^{\alpha}) - \gamma_j]\}.$$

If the unobservables are uncorrelated across consumers in the same community, then the marginal prices of consumer i' who has similar observable characteristics and lives in the same community satisfy the necessary conditions required for an instrument. Since the observable characteristics are similar, the marginal prices of the two consumers are correlated. Also, since the ε^{α} terms are independently distributed, the marginal prices for i' are not correlated with the error terms in the modified demand function (19). Basing the definition of similarity on a chosen attribute such as location of consumers in the same community may be more problematic. In this case, consumers i and i' are similar if they consume similar amounts of some attribute, say attribute n, or if $\mid z_{ij}^{d,n} - z_{i'j}^{d,n} \mid < \eta$. Given the demand functions (3), consumers similar to i are those with

$$\mid [(\Gamma_j + A)^{-1}A]_n(\alpha_{ij}^{o} + \mu_{ij}^{\alpha} + \varepsilon_{ij}^{\alpha} - \alpha_{i'j}^{o} - \mu_{i'j}^{\alpha} - \varepsilon_{i'j}^{\alpha}) \mid < \eta$$

where $[(\Gamma_j + A)^{-1}A]_n$ denotes row n of the matrix. In this case, the marginal prices of consumer i' for whom the above inequality holds are correlated with the error term in consumer i's modified demand function. Although the (unconditional) distributions of the $\varepsilon_{\bullet j}^{\alpha}\,'s$ for are independent across consumers, the distributions of $\varepsilon_{i'j}^{\alpha}$ and $\varepsilon_{ij}^{\alpha}$ conditional upon the above inequality are not.

3.4 Matching of Agents

As mentioned in a previous section, estimating the hedonic bid and offer functions or the demand and supply functions does not require knowledge of the matches between consumers and firms. The relationship between the agents who trade with one another is nonetheless of interest and provides further evidence on properties of the hedonic equilibrium. To simplify the notation for this discussion, we deal with only one community. This allows us to drop the j subscript in this subsection.

Consumers and firms that choose the same attribute vectors as given by (13) and (14) trade with one another. Specifically, a consumer of type α trades with a firm of type β if

$$(\Gamma + A)^{-1}(A\alpha - \gamma) = (\Gamma + B)^{-1}(B\beta - \gamma),$$

or equivalently if

$$\alpha = R\beta + S$$

where $R = A^{-1}\ (\Gamma + A)(\Gamma - B)^{-1}B$ and $S = (A^{-1} - RB^{-1})\gamma$. In the simplest case, with the matrices A and B diagonal and with the no correlation in the distribution of the consumer characteristics α or in the firm characteristics β, the match between consumer and firm characteristics has the simple and intuitive form

$$\alpha_n = -\frac{\sigma_{\alpha_n}}{\sigma_{\beta_n}}\beta_n + (\overline{\alpha_n} + \frac{\sigma_{\alpha_n}}{\sigma_{\beta_n}}\overline{\beta_n})$$

for $n = 1, ..., N$. Consumers who value an attribute highly are matched with firms producing the attribute at low cost. A firm with the mean value of β characteristic is matched to a consumer with the mean value of the corresponding α characteristic. Similarly, a firm with a β value one standard deviation above the mean for firm characteristic n is matched with a consumer having an α value one standard deviation below the mean for this characteristic. With the normal distributions for α and β, it is straightforward to check that the equal "numbers" of firms and consumers are matched.

Since researchers do not observe the α and β but instead only α^o and β^o, one feature of the relationship between the matched firms and consumers involves the unobservables which are the source of the error terms in the bid/offer and supply/demand equations. Consider a particular value for the firm characteristic, say $\widehat{\beta}$. Firms of this type are all matched with consumers of type $\widehat{\alpha}$ $(= R\widehat{\beta} + S)$. The value $\widehat{\beta}$ is generated by any β^o, μ^{β}, and ε^{β} for which $\widehat{\beta} = \beta^o + \mu^{\beta} + \varepsilon^{\beta}$. Obviously many different combinations could produce $\widehat{\beta}$. Similarly,

$\widehat{\alpha}$ $(= R\widehat{\beta} + S)$ could be produced by many combinations of α^o, μ^α, and ε^α. The market matches the $\widehat{\beta}$ type firms with the $\widehat{\alpha}$ type consumers but does not determine uniquely how the β^o, μ^β, and ε^β combinations yielding $\widehat{\beta}$ align with the α^o, μ^α, and ε^α combinations yielding $\widehat{\alpha}$. It is consistent with the theory for the $\widehat{\alpha}$ type consumers with the highest values of the unobservables to trade with the $\widehat{\beta}$ type firms with the highest values of the unobservables. It is also consistent for there to be completely reverse sorting in which the $\widehat{\alpha}$ type consumers with the highest values of the unobservables purchase from the $\widehat{\beta}$ type firms with the lowest values of the unobservables.

Suppose now that we are estimating the bid and offer functions as given by (13) and (14) and that we have data for matched pairs of firms and consumers who trade with one another. The theory implies nothing about whether the error terms in the bid and offer equations are positively correlated, negatively correlated, or uncorrelated. If we estimate the bid and offer equations as a system, the method should be sufficiently general to allow for any pattern of correlation. In this respect, the outcome is similar to the usual demand and supply model with homogenous goods. In the homogeneous goods model, the market matches the entire group of consumers who choose to purchase the good with the entire group of firms who choose to sell it at the market determined price. The market equilibrium determines a price such that the number of units demanded equals the number of units supplied but does not determine which consumers trade with which firms. A consumer with a positive value for the error terms appearing the demand equation (15) may be matched with a firm (or firms) having either a positive or negative value for its error term in the individual supply equation (16). The only difference for the hedonic model is the matching of consumers and firms at each attribute value rather than just the matching of the total group of consumers and firms who want to purchase the homogeneous good.

Figure 1 provides an example of the matching process for the case with only one product attribute or $N = 1$. In order to have a two dimensional graph we assume that the correlated error terms are identically zero. Further, to make the calculations simple, we construct the graphs for the case in which the means of the observables $\overline{a_j}$ and $\overline{\beta_j}$ are both zero and the variances of the distributions of α^o, $\beta^o, \varepsilon^\alpha$, and ε^β are all identical. The graph to the left (right) shows the coordinates for the consumer (firm) observable and unobservable characteristic combinations Given the assumed parameter values, firms for which $\beta^o + \varepsilon^\beta = \widehat{\beta}$ are matched with consumers for which $\alpha^o + \varepsilon^\alpha = -\widehat{\beta}$. For example, consumers on line L_1^c are matched with firms on line L_1^f. For all points on L_1^c, the sum $\alpha^o + \varepsilon^\alpha$ is 1 and for all points on L_1^f, the sum $\beta^o + \varepsilon^\beta$ is -1. The model determines this match.

The model does not determine which combinations on L_1^c are paired with which combinations on L_1^f. One possibility is that the consumer at A (B) is paired with the firm at a (b). Alternatively, the pairings might be reversed or random. The former possibility is given by a rule pairing consumer $(\alpha^o, \varepsilon^\alpha)$ with firm $(\beta^o, \varepsilon^\beta) = (-\alpha^o, -\varepsilon^\alpha)$. This pairing rule satisfies the market match and implies a negative correlation between the unobservables of consumers and firms that trade with one another. With the observables and unobservables independently distributed in the population, there is no correlation between unobservables of a firm (consumer) and observables of a consumer (firm) that trade with one another. The second possibility is given by a rule pairing consumer $(\alpha^o, \varepsilon^\alpha)$ with firm $(\beta^o, \varepsilon^\beta) = (-\varepsilon^\alpha, -\alpha^o)$. This pairing rule also satisfies the market match, In this case, the unobservables for firms and consumers that trade with one another are uncorrelated but the observable firm (consumer) characteristics are correlated with the unobservable consumer (firm) characteristics. With random pairing, the expected characteristics of the firm with which consumer $(\alpha^o, \varepsilon^\alpha)$ trades are $(\frac{-\alpha^o - \varepsilon^\alpha}{2}, \frac{-\alpha^o - \varepsilon^\alpha}{2})$. We expect the observable and unobservable characteristics of consumers to be correlated with both the observable and unobservable characteristics of the firms.

With the first pairing rule, the observed firm characteristics could be used as an instrument in estimating the bid function. Since the observed characteristics are correlated, the observed firm characteristics are correlated with the attributes chosen by a consumer. The observed firm characteristic is not, however, correlated with the bid function error term which is determined by ε^α. With the reverse or random pairing these conditions do not hold, and the observed firm characteristics cannot be used as an instrument in estimating the bid function. Except for the nongeneric case with the first type of pairing, the characteristics of the trading partner cannot be used as instruments

3.5 Bid Functions for Non-Quasilinear Preferences

If preferences are not quasilinear, then it may be more difficult to obtain closed form solutions for the bid functions and for the equilibrium hedonic price function but the essential issues involving identification of the bid functions and the possible instruments are the same. Let $U(z, y)$ denote the utility function for the general case. A consumer maximizes utility subject to $P(z) + y = I$, and the optimal consumption bundle depends upon the parameters of the hedonic price function, the consumer's observed and unobserved taste and preference parameters, and the consumer's income I.

At an interior optimum for consumer i in community j, the marginal rate of substitution between each attribute and the expenditure on all other goods equals the marginal price of the attribute, or

$$P^j_{z_n}(z_{ij}) = \frac{U_{z_n}(z^d_{ij}, y_{ij})}{U_y(z^d_{ij}, y_{ij})}$$

for $n = 1, ..., N$. The bid function is the right-hand side of the above equality. Two of the arguments of the bid function are the same as for a model with quasilinear preferences, namely the commodity attributes (z^d_{ij}) and the the vector of consumer characteristics. With preferences that are not quasilinear, the expenditure on other goods also appears as an argument of the bid function. In brief, the vector of N bid functions may be denoted $b(z_{ij}, y_{ij}, \alpha^o_{ij}.\mu^\alpha_{ij}, \varepsilon^\alpha_{ij})$. As for the case of quasilinear preferences, the error terms in the bid functions depend on the unobserved consumer characteristics (μ^α_j and ε^α_{ij}).[6] The bid functions now include one additional right-hand side variable that is correlated with the error terms (y_{ij}), but income and wealth now serve as an additional instruments The other issues involving the need for instrumental variables methods are the same as for quasilinear preferences.

4. Estimation of the Hedonic Price Function

Although the potential endogeneity concerns involved in estimation of the hedonic demand and supply functions are exactly identical to those for the classical model, the estimation of the hedonic price function gives rise to new issues since the hedonic price function is not directly observable. For notational simplicity in this section, we assume that the means of the consumer and firm characteristics differ across communities but that the variances do not. Allowing the variances to differ considerably complicates the notation in this section but give rise to any new endogeneity issues. With these simplifications, the parameter matrix Γ as given by (6) in the hedonic price function does not vary across communities. Given the assumed utility and cost functions, the equilibrium price function in community j (12) is

$$P^j(z) = (M^\beta\ \bar{\beta}_j + M^\alpha\ \bar{\alpha}_j)'z + \frac{z'\Gamma z}{2}$$

where M^α and M^β are given by (9) and (10).

In estimating hedonic price functions, it is also common to include a stochastic term (v) arising from errors in measurement or other types of specification errors In the simplest case this error term is uncorrelated with any of the variables included in the hedonic price equation, and we initially make this assumption. Writing the consumer and firm characteristics in terms of the three components of α and β and including this stochastic term, the hedonic price function for community j is

$$P^j(z) = (M^\alpha(\bar{\alpha}^o_j + \bar{\mu}^\alpha_j + \bar{\varepsilon}^\alpha_j) + M_\beta(\bar{\beta}^o_j + \bar{\mu}^\beta_j + \bar{\varepsilon}^\beta_j))'z + \frac{z'\Gamma z}{2} + \nu.$$

In a large community, the means of the purely idiosyncratic error terms ($\bar{\varepsilon}^\alpha_j$ and $\bar{\varepsilon}^\beta_j$) should be negligible and are assumed to be zero in the remainder of this discussion. Keep in mind that although researchers cannot observe the means $\bar{\mu}^\alpha_j$ and $\bar{\mu}^\beta_j$, the participants in each market know the slopes of the equilibrium price function for their market. The estimation of the hedonic price function must therefore reflect the differences that are attributable to both observable and unobservable community characteristics. We do not assume that the consumers and firms understand how the characteristics of their communities affect the $EHPF$

[6] Whether they enter linearly or nonlinearly depends upon the functional form of the bid function.

(equilibrium hedonic price function) but we do make the standard assumption that the each consumer and firm knows the function for his or her own community.

Suppose then that we have data on the means of the observable community characteristics and on the prices and attributes. There is no particular advantage to having matched data for consumers and firms apart from the possibly larger data set but we will initially use notation that allows for that possibility. Suppose then that the price and attribute data consist of matched pairs (p_{ikj},z_{ikj}) with consumer i purchasing from firm k in community j. With matched data, the equation to be estimated is

$$p_{ikj} = (M^{\alpha}(\bar{\alpha}_j^o + \bar{\mu}_j^{\alpha}) + M^{\beta}(\bar{\beta}_j^o + \bar{\mu}_j^{\beta}))' z_{ikj} + \frac{z_{ikj}^{'} \Gamma z_{ikj}}{2} + \nu_{ijk}. \tag{20}$$

The above equilibrium hedonic price equation is market solution for a model without government regulations or special geographic configurations. Extending the model to include such features (e.g., minimum child-care staffing regulations; proximity to ports or natural recreational areas) could affect the entire equilibrium hedonic price function. The precise nature of the new price function would depend upon the exact regulation or geographic feature and on the distribution of consumer and firm characteristics within the community, Although there might be no closed-form solution to the model, the equilibrium hedonic price function would include measures of the regulation or geographic features as explanatory variables in addition to the product attributes. If the regulation takes form of minimum attribute standards, then the hedonic price function would not be observed for the unallowed combinations and would differ from the unrestricted price function on the allowed combinations.

If the regulations or geographic features affect consumers and firms only through the equilibrium price function, then the measures of the regulation would appear in the hedonic price function but not as explanatory variables in the bid and offer functions. In this case, the regulatory variables or geographic features might be used as instruments in estimating the bid and offer equations unless the regulations are correlated with the unobserved consumer or firm characteristics. Diamond and Smith (1985) suggest that geographical features such as the presence of a sea port might be used to create instruments. Such features increase the marginal product of labor, ceteris paribus; higher housing prices then equalize real incomes across communities. Some features of the climate might play the same role; others, such as the severity of the winter, could directly affect the demand for some housing attributes and thus enter not only the hedonic price function but also and the bid and offer functions.

4.1 Uncorrelated Unobservable Characteristics

The alternatives for estimating the equation depend upon whether or not it is assumed that the unobservables μ_{ij}^{α} and μ_{ij}^{β}, which are correlated across consumers and firms in a community, are nonzero. If they are zero, then the above price equation simplifies to

$$p_{ikj} = (M^{\alpha}\bar{\alpha}_j^o + M^{\beta}\bar{\beta}_j^o)' z_{ikj} + \frac{z_{ikj}^{'} \Gamma z_{ikj}}{2} + \nu_{ijk}.$$

There are two possible methods for estimating the price equation in a way that reflects the differences in the equilibrium price function faced by agents in different communities. The first is to estimate separate slope terms for each community, or to estimate

$$p_{ikj} = \gamma_j' z_{ikj} + \frac{z_{ikj}^{'} \Gamma z_{ikj}}{2} + \nu_{ijk} \tag{21}$$

where γ_j denotes the slope coefficients for community j. In empirical applications it is commonly assumed that only a subset of the coefficients differ across communities, and practical considerations related to the number of communities and the size of the data set often affect the number of coefficients that are allowed to differ.

The second alternative is to capture the differences in the $EHPF$ through the observable community characteristics and to estimate the matrices M_{α}, M_{β}, and Γ with the restriction that Γ is symmetric. There are trade-offs in the data requirements for the two methods. The first requires a relatively large number of observations from each community whereas the second requires data on the mean community characteristics

$\bar{\alpha}_j^o$ and $\bar{\beta}_j^o$. The marginal prices computed using either method reflect the price differences faced by consumers in different communities.

4.2 Correlated Unobservable Characteristics

If it is not reasonable to assume *a priori* that the unobservables μ_{ij}^{α} and μ_{ij}^{β} are zero, then only the first approach described above is completely satisfactory. In order to understand the drawbacks to the second approach in this case, rewrite (20) to group the error components together

$$p_{ikj} = (M^{\alpha}\bar{\alpha}_j^o + M^{\beta}\ \bar{\beta}_j^o)' z_{ikj} + \frac{z_{ikj}^{'}\Gamma z_{ikj}}{2} + \eta_j' z_{ikj} + \nu_{ijk} \tag{22}$$

where the error term η_j equals $(M^{\alpha}\bar{\mu}_j^{\alpha} + M^{\beta}\bar{\mu}_j^{\beta})$. There are two complications, one of which is technical and surmountable and the second of which is conceptual and unavoidable. The technical difficulty comes from the presence of the two error terms (η_j and ν_{ijk}). The term η_j, enters non additively. In addition, it is correlated with the explanatory variable z_{ikj} since the parameters of the equilibrium hedonic price function and hence the chosen attributes depend on $\bar{\mu}_j^{\alpha}$ and $\bar{\mu}_j^{\beta}$. Instrumental variables estimation is feasible using the observed individual consumer or firm characteristics as instruments. The individual characteristics do not appear in the price equation and are not correlated with the error terms and yet affect the individual attribute choices. In a large and diverse community the observed characteristics of the individual agents would be quite different from the averages of the observed characteristics ($\bar{\alpha}_j^o$ and $\bar{\beta}_j^o$) which do appear in the price equation.

The conceptual difficulty is that the resulting estimates do not completely capture the differences in the marginal prices across communities as perceived by the consumers and firms. The agents perceive the actual marginal prices which are

$$[M^{\alpha}\ \bar{\alpha}_j^o + M^{\beta}\ \bar{\beta}_j^o + \Gamma z] + M^{\alpha}\bar{\mu}_j^{\alpha} + M^{\beta}\ \bar{\mu}_j^{\beta}$$

for attribute vector z. At best, the estimates of the marginal prices arising from the second method equal the bracketed term, which is the average of the marginal prices for communities with the same mean observable characteristics as community j.

4.3 Hedonic Price Function Error Term

Thus far we have assumed that the error term, $\nu_{ijk,}$ in the hedonic price function is uncorrelated with the commodity attributes, but there are common circumstances in which this would not be the case. Suppose, for example, that the error term incorporates factors observed by the consumers and firms but not the researcher. Further, assume non-quasilinear preferences in which income affects the consumption of the product attributes. In these circumstances, the value of the error term in the hedonic price function affects the attributes demanded. Indeed, it is far easier to provide common-sense explanations for why the error term in the hedonic price function would be correlated with the attributes than to provide explanations for why it would not be correlated.[7].

If some community specific parameters are estimated, then either the individual consumer and firm characteristics or community characteristics are feasible instruments for estimating the hedonic price function. These characteristics do not appear in the price function and are generally assumed to be uncorrelated with the error term in the price function (e.g., Kahn and Lang, 1988 or Wooldridge, 1996). The individual characteristics directly affect the chosen product attributes, and their means affect the chosen attributes through their effects on the parameters of the hedonic price function. If the hedonic price function error term, $\nu_{ijk,}$, is uncorrelated across consumers and firms in a community, then community binaries and the

[7]Chay and Greenstone (2000) consider a related issue in their study of air quality and housing prices. Using county level data, they estimate median housing prices as dependent upon county characteristics such as population density and per capita county expenditures, which are treated as exogenous, and upon average air quality. In contrast to previous studies, they allow for the possibility that air quality is correlated with the price function error term. The correlation is attributable to local economic shocks that reduce overall economic activity hence lowering housing prices and decreasing pollution. In the hedonic model for private goods, which is the model discussed in this paper, the source of the correlation between the attributes and the error term in the hedonic price equation is slightly different. Rather than having a third factor that affects both the attribute (air quality) and the price, the error term in the price equation directly affects the attributes chosen by a consumer.

attributes chosen by similar consumers can also be used as instruments. If $\bar{\mu}_j^{\alpha}$ and $\bar{\mu}_j^{\beta}$ are identically zero and the differences in the price function across communities are captured by including the means of the consumer and firm characteristics in the hedonic price function, then only the individual characteristics can be used as instruments

4.4 Joint Estimation

The bid and offer equations along with the $EHPF$ could be estimated as a system consisting of (17), (18), and (21). The conditions under which instrumental variables estimation is required are the same as for the equation-by-equation estimation, and the feasible instruments are also the same. Wooldridge (1996) describes the efficient GMM estimator for such a system. Given the potential correlation between the observed consumer (firm) characteristics and the errors of the offer (bid) equations as described in the section on matching, 3SLS is inefficient.[8]

Regardless of the assumptions about the correlation of the error term in the hedonic price function with the attributes, the estimation of the hedonic price function has the somewhat peculiar feature that the right-hand side variables z_{ikj} are functions of the estimated parameters. Consider even the simplest case in which $\bar{\mu}_j^{\alpha}$and $\bar{\mu}_j^{\beta}$ are zero and the hedonic price function is given by (21). The attributes demanded by consumer i in community j (z_{ij}^d) depend on the parameters γ_j and Γ of the hedonic price function which are the parameters being estimated. Similarly, the attributes supplied by firm k in community j depend on the estimated parameters.

The parameters of such a model are not necessarily identified. As shown by Brown and Rosen (1982), recovering the parameters of a hedonic price function requires either (i) *a priori* restrictions on the functional forms of the bid/offer functions relative to the hedonic price functions or (ii) data from more than one community. Further, Mendelsohn (1985) considers the class of polynomial bid, offer and hedonic price functions and determines the restrictions that must be placed on the highest power of the bid and offer functions relative to those of the price function in order to identify the functions using data from only one community. McConnell and Phipps (1987) extend this analysis to other functional forms for the hedonic price function and the bid functions and discuss conditions under which the parameters of the utility function can be recovered.

The importance of the functional forms has also arisen in recent work in which community choice is endogenous. Although we allow for correlation in the unobservables across consumers/firms in a community, we take the distribution of consumers/firms in communities as given. Diamond and Smith (1985) and others have noted that the usual hedonic model does not deal with the simultaneous attainment of equilibrium in the market for attributes and the formation of communities. Nesheim (2000) has addressed this issue by developing a model in which the distribution of consumers and the hedonic price function are jointly determined. Each consumer cares about the characteristics of others living in the community, and in equilibrium the consumers may group by characteristics. Nesheim determines functional forms for which the model is identified.

5. Summary

Tables 1 summarizes the conditions under which instrumental variables estimation is required in estimating the demand/bid functions and lists potential instruments. The conditions for the supply-type functions are analogous. Table 2 summarizes the same conditions for the hedonic price function. Although we have not explicitly discussed estimation using data for only one community, this situation may be treated as a special case of no correlation in the consumer and firm unobservables within a community and no correlation in the hedonic price function errors within a community. Table 2 covers this case for the estimation of the hedonic price function, and Table 3 summarizes the conditions under which instrumental variables estimation is required in estimating the demand/bid functions.

[8] The price variables that appear on the right-hand side of the (modified) demand/supply demand functions depend upon the estimated hedonic price function parameters. Even if the functions are not estimated as a system, the estimation and the computation of the standard errors for the coefficients of the demand/supply functions must be consistent with using estimated parameters (or functions of them) as right-hand side variables.

6. References

P. Anglin and R. Gencay (1996), "Semiparametric Estimation of a Hedonic Price Function," *Journal of Applied Econometrics 11, pp. 633-648.*

T. Bartik (1987), "The Estimation of Demand Parameters in Hedonic Price Models," *The Journal of Polictical Economy* 95, pp. 81-88..

E. Berndt and M. Showalter (1993), "An Empirical Investigation of the Box-Cox Model and a Nonlinear Least Squares Alternative," *Econometric Reviews 12,* pp. *65-102.*

K. Chay and M. Greenstone (2000), "Does Air Quality Matter? Evidence from the Housing Market," Working Paper.

P. Cheshire and S. Sheppard (1998), "Estimating the Demand for Housing, Land, and Neighborhood Characteristics," *Oxford Bulletin of Economics and Statistics.*

D. B. Diamond, Jr. and B. A. Smith (1985), "Simultaneity in the Market for Housing Characteristics", *Journal of Urban Economics* 17, 280-292.

D. Epple (1987), "Hedonic Prices and Implicit Markets; Estimating Demand and Supply Functions for Differentiated Products," *Journal of Political Economy* 95, pp. 59-80.

S. Kahn and K. Lang (1988), "Efficient Estimation of Structural Hedonic Systems", *International Economic Review 29,* pp. 157-166.

K. E. McConnell and T. T. Phipps (1987), "Identification of Preference Parameters in Hedonic Models: Consumer Demands with Nonlinear Budgets", *Journal of Urban Economics* 22, 35-72.

R. Mendelsohn (1985), "Identifying Structural Equations with Single Market Data", *Review of Economics and Statistics* 29, pp. *525-529.*

M. Murray (1983), "Mythical Demands and Mythical Supplies for Proper Estimation of Rosen's Hedonic Price Model," *Journal of Urban Economics* 14, pp. 327-337.

L. Nesheim (2000), "Equilibrium Sorting of Heterogeneous Consumers Across Locations: Theory and Empirical Implications," Working Paper 2000.

S. Rosen (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy* 82, pp. 34-55.

Stephen Sheppard (1999), "Hedonic Analysis of Housing Markets" *Handbook of Regional and Urban Economics* Volume 3, Chapter 41. 1999.

J. M. Wooldridge (1996), "Estimating Systems of Equations with Different Instruments for Different Equations" *Journal of Econometrics* 74, pp. 387

S.H. Yoo (2001), "A Robust Estimation of Hedonic Price Models: Least Absolute Deviations Estimation, " *Applied-Economics-Letters* 8, pp. 55-58.

Table 1: Estimation of Demand/Offer Functions for the Hedonic Model

Functions Estimated **Assumption about Unobservables**	**Dependent Variables**	**Independent Variables in Addition to Consumer Characteristics**	**IV Estimation Instruments in Addition to the Included Arguments** (along with interactions of the variables listed)
Demand functions Unobservables not correlated across consumers within a community	Attributes demanded	Estimated parameters of the hedonic price function	IV unnecessary
Demand functions Unobservables correlated across consumers within a community	Attributes demanded	Estimated parameters of the hedonic price function	Means and other moments of community characteristics
Bid functions Unobservables not correlated across consumers within a community	Estimated marginal prices	Attributes demanded	Means and other moments of community characteristics Community binaries Attributes chosen by consumers with similar characteristics
Bid functions Unobservables correlated across consumers within a community	Estimated marginal prices	Attributes demanded	Means and other moments of community characteristics
Modified demand functions Unobservables not correlated across consumers within a community	Attributes demanded	Estimated marginal prices	Same as for bid function with unobservables not correlated
Modified demand functions Unobservables not correlated across consumers within a community	Attributes demanded	Estimated marginal prices	Same as for bid function with unobservables correlated across consumers within a community

Table 2: Estimation of the Hedonic Price Function

A. Community Differences Captured Parametrically by Including Observed Community Characteristics As Regressors

Assumption about the Price Function Error Term	**Assumption about the Consumer/Firm Unboservables**	**Arguments of the Hedonic Price Function in Addition to Community Characteristics that Affect the Price But Not Utility or Cost Directly**	**IV Estimation Instruments in Addition to the Included Arguments** (along with interactions of the variables listed)
Uncorrelated w/ the attributes	Uncorrelated	Attributes Community characteristics	IV unnecessary
Correlated w/ the attributes but not across agents within a community	Uncorrelated	Attributes Community characteristics	Consumer or firm characteristics Community binaries (assuming not colinear w/ the community characteristics)
Correlated w/ the attributes and correlated across agents within a community	Uncorrelated	Attributes Community characteristics	Consumer or firm characteristics

B. Community Differences Captured by Estimating Some Community-Specific Coefficients			
Assumption about the Price Function Error Term	**Assumption about the Consumer/Firm Unboservables**	**Arguments of the Hedonic Price Function in Addition to Community Characteristics that Affect the Price But Not Utility or Cost Directly**	**IV Estimation Instruments in Addition to the Included Arguments** (along with interactions of the variables listed)
Uncorrelated w/ the attributes	Uncorrelated or correlated within a community	Attributes	IV unnecessary
Correlated w/ the attributes but not across agents within a community	Uncorrelated or correlated within a community	Attributes	Individual consumer or firm characteristics Community characteristics or dummies Attributes chosen by consumers living in the same community and having similar characteristics
Correlated w/ the attributes and correlated across agents within a community	Uncorrelated or correlated within a community	Attributes	Individual consumer or firm characteristics Community characteristics

C. Community Differences Captured by Estimating Separate Hedonic Price Functions for Each Community
(or only one community in data set)

Assumption about the Price Function Error Term	Arguments of the Hedonic Price Function	IV Estimation Instruments in Addition to the Included Arguments (along with interactions of the variables listed)
Uncorrelated w/ the attributes	Attributes	IV unnecessary
Correlated w/ the attributes but not across agents within a community	Attributes	Individual consumer or firm characteristics Community characteristics or dummies Attributes chosen by consumers living in the same community and having similar characteristics

Table 3: Estimation of Demand/Offer Functions with Data for Only One Community

(Assuming no correlation in the unobservables across consumers or firms within a community and functional forms that provide the demand/offer functions)

Functions Estimated **Assumption about Unobservables**	**Dependent Variables**	**Independent Variables in Addition to Consumer Characteristics**	**IV Estimation Instruments in Addition to the Included Arguments** (along with interactions of the variables listed)
Demand functions	Attributes demanded	Estimated parameters of the hedonic price function	IV unnecessary
Bid functions	Estimated marginal prices	Attributes demanded	Attributes chosen by consumers with similar observable characteristics
Modified demand functions	Attributes demanded	Estimated marginal prices	Attributes chosen by consumers with similar observable characteristics

Figure 1: Matching of Consumers and Firms

Panel a: Distribution of Consumer Characteristics

Panel b: Distribution of Firm Characteristics